OM protein - protein search, using sw model

Run on: April 27, 2006, 15:09:06 ; Search time 220.235 Seconds
(without alignments)
103.742 Million cell updates/sec

Title: US-09-819-144A-2
Perfect score: 268
Sequence: 1 MAAGPRPGAPCRAGAPTIVL.....SSPARATLGKPLVLNDYSLN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	268	100.0	52	2	AAY39293	Aay39293 CSA-1 cho
2	77	28.7	209	7	ABM87430	Abm87430 Rice abio
3	74	27.6	16	2	AAY39295	Aay39295 Polypepti
4	73.5	27.4	337	7	ADC32784	Adc32784 Human nov
5	73.5	27.4	675	9	AEA52600	Aea52600 Human Rab
6	73.5	27.4	791	5	ABP65089	Abp65089 Hypoxia-r
7	73.5	27.4	791	8	ADN03640	Adn03640 Antipsori
8	73.5	27.4	791	8	ADP22972	Adp22972 PRO polyp
9	73.5	27.4	863	6	ABP98856	Abp98856 Human str
10	73.5	27.4	863	7	ADC31071	Adc31071 Human nov
11	73.5	27.4	863	7	ADE48296	Ade48296 Human MIC
12	73.5	27.4	863	7	ADE48308	Ade48308 Human MIC
13	73.5	27.4	863	8	ABM82355	Abm82355 Tumour-as
14	72.5	27.1	173	7	ABO76575	Abo76575 Pseudomon
15	71.5	26.7	76	3	AAG28488	Aag28488 Zea mays
16	71	26.5	19938	6	ABP76682	Abp76682 Streptomy
17	69.5	25.9	210	4	AAU48625	Aau48625 Propionib
18	69.5	25.9	210	6	ABM45144	Abm45144 Propionib
19	68	25.4	315	5	ABU05772	Abu05772 M. tuberc
20	68	25.4	315	8	ABM79606	Abm79606 M tubercu
21	67.5	25.2	135	4	AAO01165	Aao01165 Human pol
22	67.5	25.2	137	7	ABO81794	Abo81794 Pseudomon
23	67	25.0	1321	9	ADV97797	Adv97797 Murine pr
24	66	24.6	228	7	ABO79635	Abo79635 Pseudomon
25	65.5	24.4	923	7	ADD18710	Add18710 Human dis
26	65.5	24.4	923	8	ADO19848	Ado19848 Human PRO
27	65.5	24.4	924	8	ADU06294	Adu06294 Novel bro
28	65	24.3	166	7	ABO68030	Abo68030 Pseudomon
29	65	24.3	205	7	ABO82421	Abo82421 Pseudomon
30	65	24.3	345	7	ABO77022	Abo77022 Pseudomon
31	64.5	24.1	146	7	ABO77184	Abo77184 Pseudomon
32	64.5	24.1	149	7	ABO82120	Abo82120 Pseudomon
33	64.5	24.1	309	4	ABG12975	Abg12975 Novel hum
34	64.5	24.1	341	8	ADK71091	Adk71091 Human MP2
35	64	23.9	137	3	AAG28496	Aag28496 Zea mays
36	63.5	23.7	223	8	ADY06811	Ady06811 Plant ful
37	63.5	23.7	377	9	ADY85271	Ady85271 Human ort
38	63.5	23.7	478	7	AAO30817	Aao30817 Human cel
39	63.5	23.7	721	7	ABO79547	Abo79547 Pseudomon
40	63	23.5	53	5	ABP01580	Abp01580 Human ORF
41	63	23.5	128	7	ADB64958	Adb64958 Human pro
42	63	23.5	128	7	ADM03909	Adm03909 Human pro
43	63	23.5	613	7	ABO71209	Abo71209 Pseudomon
44	63	23.5	1129	7	ABO73584	Abo73584 Pseudomon
45	62.5	23.3	85	4	AAU21041	Aau21041 Human nov

ALIGNMENTS

RESULT 1
AAY39293
ID AAY39293 standard; protein; 52 AA.
XX
AC AAY39293;
XX
DT 26-NOV-1999 (first entry)
XX
DE CSA-1 chondrosarcoma associated protein-1.
XX
KW Chondrosarcoma associated protein-1; CSA-1; bone malignancy; diagnosis;
KW rheumatoid arthritis; inflammatory arthropathy; tumour;
KW cartilage associated polypeptide; CAA-1.
XX
OS Homo sapiens.
XX
PN WO9946382-A1.
XX
PD 16-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US005348.
XX
PR 13-MAR-1998; 98US-00042225.
XX
PA (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.
XX
PI Terek RM;
XX
DR WPI; 1999-551411/46.
DR N-PSDB; AAZ06793.
XX
PT New gene encoding a chondrosarcoma associated (CSA) polypeptide, useful
PT for diagnosing bone malignancy.
XX
PS Claim 10; Page 18; 47pp; English.
XX
CC This is the human chondrosarcoma associated protein-1 (CSA-1) amino acid
CC sequence. Chondrosarcoma is the second most common form of bone
CC malignancy and occurs in late adulthood and old age. CSA-1 is expressed
CC in a tumour cell line and also in some high grade chondrosarcoma, but not
CC in normal cartilage, or low or intermediate grade tumours. The CSA-1
CC polynucleotide and polypeptide can be used in methods and compositions
CC for evaluating appropriate treatment and treatment effectiveness of
CC malignancies associated with expression of CSA-1. CSA-1 polynucleotide
CC can be used as a probe to classify cells in terms of their level of CSA-1
CC expression or as primers for diagnostic PCR analysis in which mutations
CC and allelic variation of CSA-1 can be detected. Transgenic animals

```
CC   containing human CSA-1 or with a null mutation can serve as models for
CC   chondrosarcoma. Methods of treating undesired inflammation such as that
CC   associated with rheumatoid arthritis and other inflammatory arthopathies
CC   is carried out by administering an effective amount of a cartilage
CC   associated (CAA) polypeptide such as CAA-1 AAY39294
XX
SQ   Sequence 52 AA;

  Query Match             100.0%;  Score 268;  DB 2;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-27;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSLN 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSLN 52


RESULT 2
ABM87430
ID   ABM87430 standard; protein; 209 AA.
XX
AC   ABM87430;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Rice abiotic stress responsive polypeptide SEQ ID NO:5676.
XX
KW   abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
XX
OS   Oryza sativa.
XX
PN   WO2003008540-A2.
XX
PD   30-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019668.
XX
PR   22-JUN-2001; 2001US-0300112P.
PR   24-AUG-2001; 2001US-0314662P.
PR   26-SEP-2001; 2001US-0325277P.
PR   21-NOV-2001; 2001US-0332132P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Kreps J,  Briggs SP,  Cooper B,  Glazebrook J,  Goff SA,  Katagiri F;
PI   Moughamer T,  Provart N,  Ricke D,  Zhu T;
XX
DR   WPI; 2003-248011/24.
XX
PT   New stress-responsive nucleic acid, useful for altering the
PT   responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT   stress, salt stress or osmotic stress.
XX
PS   Claim 1; SEQ ID NO 5676; 89pp; English.
XX
CC   The invention relates to novel abiotic stress responsive polynucleotides
CC   and polypeptides. Also disclosed are vectors, expression cassettes, host
CC   cells, and plants containing such polynucleotides. Also disclosed are
CC   methods for using the polynucleotides and polypeptides to alter the
CC   responsiveness of a plant to abiotic stress. The invention is useful in
CC   agriculture. The nucleic acid is useful for determining whether a test
CC   plant has been exposed to an abiotic stress condition. It is also useful
CC   for selecting an agent that alters abiotic stress regulated
CC   polynucleotide expression in a plant cell, and to identify a homolog or
CC   ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC   molecule and the polypeptide encoded by it are useful in altering the
CC   responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC   stress, osmotic stress or any of their combinations. The present sequence
CC   is used in the exemplification of the invention
XX
SQ   Sequence 209 AA;

  Query Match             28.7%;  Score 77;  DB 7;  Length 209;
  Best Local Similarity   35.1%;  Pred. No. 0.13;
  Matches   20;  Conservative   11;  Mismatches   16;  Indels   10;  Gaps    3;

Qy          1 MAAGPRPGAP------CRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSL 51
              :|  | ||||      |  | ||:::| |    :|:|::| : | |   |::   ||
Db        150 LAPAP-PGAPSPGITICSGGPPTVIMTPG---VISYGAASRSSANLSSSLLVAMVSL 202


RESULT 3
AAY39295
ID   AAY39295 standard; peptide; 16 AA.
XX
AC   AAY39295;
XX
DT   26-NOV-1999  (first entry)
XX
DE   Polypeptide used to generate CSA-1-reactive polyclonal antibodies.
XX
KW   Chondrosarcoma associated protein-1; CSA-1; bone malignancy; diagnosis;
KW   rheumatoid arthritis; inflammatory arthropathy; tumour; immunogen;
KW   cartilage associated polypeptide; CAA-1; polyclonal antibody generation.
XX
OS   Homo sapiens.
XX
PN   WO9946382-A1.
XX
PD   16-SEP-1999.
XX
PF   12-MAR-1999;   99WO-US005348.
XX
PR   13-MAR-1998;   98US-00042225.
XX
PA   (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.
XX
PI   Terek RM;
XX
DR   WPI; 1999-551411/46.
XX
PT   New gene encoding a chondrosarcoma associated (CSA) polypeptide, useful
PT   for diagnosing bone malignancy.
XX
PS   Example 3; Page 24; 47pp; English.
XX
CC   This sequence is a peptide used as an immunogen to stimulate the
CC   production of human chondrosarcoma associated protein-1 (CSA-1:AAY39293)
CC   reactive polyclonal antibodies. Chondrosarcoma is the second most common
CC   form of bone malignancy and occurs in late adulthood and old age. CSA-1
CC   is expressed in a tumour cell line and also in some high grade
CC   chondrosarcoma, but not in normal cartilage, or low or intermediate grade
CC   tumours. The CSA-1 polynucleotide and polypeptide can be used in methods
CC   and compositions for evaluating appropriate treatment and treatment
CC   effectiveness of malignancies associated with expression of CSA-1. CSA-1
CC   polynucleotide can be used as a probe to classify cells in terms of their
CC   level of CSA-1 expression or as primers for diagnostic PCR analysis in
CC   which mutations and allelic variation of CSA-1 can be detected.
CC   Transgenic animals containing human CSA-1 or with a null mutation can
CC   serve as models for chondrosarcoma. Methods of treating undesired
CC   inflammation such as that associated with rheumatoid arthritis and other
CC   inflammatory arthopathies is carried out by administering an effective
CC   amount of a cartilage associated (CAA) polypeptide such as CAA-1 AAY39294
XX
SQ   Sequence 16 AA;

  Query Match             27.6%;  Score 74;  DB 2;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.015;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         24 RRQTLSHGSSSPARA 38
              |||||||||||||||
Db          1 RRQTLSHGSSSPARA 15


RESULT 4
```

```
ADC32784
ID   ADC32784 standard; protein; 337 AA.
XX
AC   ADC32784;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human novel contig-encoded polypeptide sequence, SEQ ID NO:2866.
XX
KW   Human; diagnostic; drug screening; forensics; gene mapping;
KW   biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW   neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW   ulcers; osteoporosis; autoimmune disease; cancer;
KW   molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW   neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW   antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW   gene therapy; chromosome 22q13.1-13.2.
XX
OS   Homo sapiens.
XX
PN   WO2003029271-A2.
XX
PD   10-APR-2003.
XX
PF   24-SEP-2002; 2002WO-US030474.
XX
PR   24-SEP-2001; 2001US-0324631P.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang TY,  Zhang J,  Ren F,  Xue AJ,  Zhao QA,  Wang J,  Wehrman T;
PI   Zhou P,  Ghosh M,  Wang D,  Ma Y,  Asundi V,  Wang Z,  Weng G;
PI   Haley-Vicente D,  Drmanac RT;
XX
DR   WPI; 2003-371981/35.
DR   N-PSDB; ADC32017.
XX
PT   New polynucleotide and polypeptide useful for diagnosing, preventing or
PT   treating conditions such as neurodegenerative diseases, anemias, platelet
PT   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT   cancer.
XX
PS   Example 2; SEQ ID NO 2866; 1185pp; English.
XX
CC   The invention relates to 971 novel human cDNA sequences (ADC29919-
CC   ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC   invention also relates to nucleic acid sequences over 99% identical with
CC   the novel human cDNAs. The invention additionally encompasses expression
CC   vectors and host cells comprising a nucleic acid of the invention; the
CC   recombinant production of a polypeptide of the invention; an antibody
CC   against a polypeptide of the invention; a method of detecting
CC   polynucleotides or polypeptides of the invention; and methods of
CC   identifying a compound which binds to a polypeptide of the invention. The
CC   invention further discloses methods of peventing, treating or
CC   ameliorating a medical condition; kits comprising polynucleotide probes
CC   and/or monoclonal antibodies for carrying out the methods of the
CC   invention; methods for the identification of compounds that modulate the
CC   expression or activity of the polynucleotide and/or polypeptide; and 767
CC   contig sequences corresponding to the cDNA sequences of the invention
CC   (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC   -ADC33394). The nucleic acids and polypeptides of the invention are
CC   useful in diagnostics, drug screening, forensics, gene mapping, in the
CC   identification of mutations responsible for genetic disorders or other
CC   traits, for assessing biodiversity, and in producing many other types of
CC   data and products dependent on DNA and amino acid sequences. They are
CC   also used for treating diseases such as Parkinson's disease, Alzheimer's
CC   disease and other neurodegenerative diseases, anaemia, platelet
CC   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC   cancer. The nucleic acids may also be used as hybridisation probes or
CC   primers, and in the recombinant production of a protein. The polypeptides
CC   are also useful in generating antibodies, as molecular weight markers,
CC   and as food supplements. The present sequence represents a human contig-
CC   encoded polypeptide sequence used in an example of the invention. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 337 AA;

  Query Match             27.4%;  Score 73.5;  DB 7;  Length 337;
  Best Local Similarity   45.7%;  Pred. No. 0.68;
  Matches   16;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db         62 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 96


RESULT 5
AEA52600
ID   AEA52600 standard; protein; 675 AA.
XX
AC   AEA52600;
XX
DT   25-AUG-2005  (first entry)
XX
DE   Human Rab13-related cerebral aneurysm marker protein.
XX
KW   genetic marker; brain hemorrhage; cerebroprotective; hemostatic;
KW   cardiovascular disease; neurological disease; arteriosclerosis;
KW   antiarteriosclerotic; Rab13.
XX
OS   Homo sapiens.
XX
PN   JP2005151854-A.
XX
PD   16-JUN-2005.
XX
PF   21-NOV-2003; 2003JP-00393161.
XX
PR   21-NOV-2003; 2003JP-00393161.
XX
PA   (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH.
XX
PI   Kubota M,  Machida T,  Uchino F,  Kobayashi E,  Saeki N,  Yamaura A;
PI   Hiwasa T,  Takiguchi M;
XX
DR   WPI; 2005-410933/42.
DR   N-PSDB; AEA52601.
XX
PT   New polypeptide marker, useful for specifically detecting and diagnosing
PT   human subarachnoid hemorrhagic lesion.
XX
PS   Claim 1; SEQ ID NO 3; 14pp; Japanese.
XX
CC   The invention relates to a novel polypeptide marker for diagnosing human
CC   subarachnoid hemorrhage comprising SEQ ID No. 1-23 (odd SEQ ID numbers)
CC   as given in the specification, or their partial amino acid sequences. The
CC   polypeptide of the invention may be useful for early, rapid and efficient
CC   detection and diagnosis of human subarachnoid hemorrhagic disease and
CC   subarachnoid hemorrhagic lesion e.g. arteriosclerosis. The current
CC   sequence is that of the human Rab13-related cerebral aneurysm marker
CC   protein of the invention.
XX
SQ   Sequence 675 AA;

  Query Match             27.4%;  Score 73.5;  DB 9;  Length 675;
  Best Local Similarity   45.7%;  Pred. No. 1.6;
  Matches   16;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db        588 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 622


RESULT 6
```

```
ABP65089
ID   ABP65089 standard; protein; 791 AA.
XX
AC   ABP65089;
XX
DT   12-NOV-2002  (first entry)
XX
DE   Hypoxia-repressed protein #12.
XX
KW   Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
KW   antiinflammatory; vulnerary; gynecological; ophthalmological; vaccine;
KW   hypoxia; tumourigenesis; angiogenesis; apoptosis; cancer;
KW   ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
KW   preeclapmsia; atherosclerosis; inflammatory condition; wound healing;
KW   inflammation; erythropoiesis; hair loss; human.
XX
OS   Homo sapiens.
XX
PN   WO200246465-A2.
XX
PD   13-JUN-2002.
XX
PF   10-DEC-2001; 2001WO-GB005458.
XX
PR   08-DEC-2000; 2000GB-00030076.
PR   08-FEB-2001; 2001GB-00003156.
PR   25-OCT-2001; 2001GB-00025666.
XX
PA   (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI   White J,  Mundy CR,  Ward NR,  Krige D,  Kingsman SM,  Harris RA;
PI   Rayner WN;
XX
DR   WPI; 2002-627238/67.
XX
PT   Identifying a gene involved in disease for treating hypoxia-regulated
PT   conditions, comprises comparing the transcriptome/proteome of two cell
PT   types under different conditions and identifying a differentially
PT   regulated gene.
XX
PS   Claim 13; Page 290; 538pp; English.
XX
CC   The present invention relates to methods for identifying genes and
CC   proteins that are implicated in a specific disease or physiological
CC   condition. The method comprises comparing the transcriptome/proteome of a
CC   specialised cell type implicated in a disease or condition with that of a
CC   second specialised cell type, under two experimental conditions, and
CC   identifying a gene that is differentially regulated in the two
CC   specialised cell types under experimental conditions. ABV77873-ABV78116
CC   and ABP65061-ABP65257 were identified using the methods of the invention.
CC   The coding sequences and proteins are useful for treating a disease in a
CC   patient, for manufacture of a medicament for treating hypoxia-regulated
CC   conditions, and for regulating tumourigenesis, angiogenesis, apoptosis,
CC   biological response to hypoxia conditions, or hypoxic-associated
CC   pathology in a patient. The coding sequences and proteins are also useful
CC   for monitoring the therapeutic treatment of a disease or physiological
CC   condition, such as cancer, ischaemic conditions, reperfusion injury,
CC   retinopathy, neonatal stress, preeclapmsia, atherosclerosis, inflammatory
CC   conditions, wound healing, inflammation, erythropoiesis or hair loss
XX
SQ   Sequence 791 AA;

  Query Match             27.4%;  Score 73.5;  DB 5;  Length 791;
  Best Local Similarity   45.7%;  Pred. No. 1.9;
  Matches   16;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db        516 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 550


RESULT 7
ADN03640
ID   ADN03640 standard; protein; 791 AA.
XX
AC   ADN03640;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic protein sequence #17.
XX
KW   antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   N-PSDB; ADN03639.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 9; SEQ ID NO 34; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polypeptides of the invention.
XX
SQ   Sequence 791 AA;

  Query Match             27.4%;  Score 73.5;  DB 8;  Length 791;
  Best Local Similarity   45.7%;  Pred. No. 1.9;
  Matches   16;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db        516 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 550


RESULT 8
ADP22972
ID   ADP22972 standard; protein; 791 AA.
XX
AC   ADP22972;
XX
DT   18-NOV-2004  (first entry)
XX
DE   PRO polypeptide SEQ ID NO:66.
XX
KW   PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW   osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW   antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX
OS   Unidentified.
XX
PN   WO2004041170-A2.
XX
PD   21-MAY-2004.
XX
PF   30-OCT-2003; 2003WO-US034312.
XX
```

PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-419628/39.
DR N-PSDB; ADP22971.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 7; SEQ ID NO 66; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
CC of the invention.
XX
SQ Sequence 791 AA;

Query Match 27.4%; Score 73.5; DB 8; Length 791;
Best Local Similarity 45.7%; Pred. No. 1.9;
Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

Qy 2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
::||:| || || :| | | : | |||||
Db 516 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 550

RESULT 9
ABP98856
ID ABP98856 standard; protein; 863 AA.
XX
AC ABP98856;
XX
DT 15-JUL-2003 (first entry)
XX
DE Human structural and cytoskeletal associated protein #47.
XX
KW Cytostatic; antiarteriosclerotic; anticonvulsant; nootropic; antianginal;
KW neuroprotective; cerebroprotective; hypotensive; cardiant; osteopathic;
KW antiinflammatory; antiarthritic; virucide; gene therapy; human; stroke;
KW structural and cytoskeleton-associated protein; SCAP; cancer; angina;
KW atherosclerosis; epilepsy; Huntington's disease; hypertension;
KW heart failure; osteoporosis; osteoarthritis.
XX
OS Homo sapiens.
XX
PN WO2003031940-A2.
XX
PD 17-APR-2003.
XX
PF 10-OCT-2002; 2002WO-US032851.
XX
PR 12-OCT-2001; 2001US-0328931P.
PR 19-OCT-2001; 2001US-0360681P.
PR 02-NOV-2001; 2001US-0343896P.
PR 09-NOV-2001; 2001US-0346308P.
PR 16-NOV-2001; 2001US-0332385P.
PR 07-DEC-2001; 2001US-0340776P.
PR 11-JAN-2002; 2002US-0347703P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Becha SD, Bhatia U, Blake JJ, Borowsky ML, Burrill JD, Chang H;
PI Chawla NK, Elliott VS, Emerling BM, Forsythe IJ, Gorvad AE;
PI Griffin JA, Hafalia AJA, Ho A, Ison CH, Kable AE, Khare R, Lal PG;
PI Lee S, Lee EA, Lee SY, Lehr-Mason PM, Li JX, Lindquist EA, Luo W;
PI Marquis JP, Ramkumar J, Richardson TW, Sprague WW, Swarnakar A;
PI Tang YT, Warren BA, Yang J, Yue H, Zebarjadian Y, Zheng W;
XX
DR WPI; 2003-403125/38.
DR N-PSDB; ACC44343.
XX
PT New human structural and cytoskeleton-associated proteins (SCAP) useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant SCAP expression e.g. cancer, osteoporosis, or epilepsy.
XX
PS Claim 1; Page 303-305; 361pp; English.
XX
CC This sequence represents a novel isolated human structural and
CC cytoskeleton-associated protein (SCAP) polypeptide. The polypeptides and
CC polynucleotides encoding them are useful in diagnosing, treating and
CC preventing diseases or conditions associated with the decreased
CC expression or over expression of SCAP, such as cell proliferative (e.g.
CC cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's
CC disease, stroke), heart (e.g. hypertension, heart failure, angina) and
CC skeletal muscle disorders (e.g. osteoporosis, osteoarthritis) or viral
CC infections. These are also useful in assessing the effects of exogenous
CC compounds on the expression of nucleic acid and amino acid sequences of
CC SCAP. The SCAP or its fragments are useful in screening compounds for
CC effectiveness as agonist or antagonist of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to or modulate the activity of the polypeptide. The
CC microarray is useful in monitoring or measuring protein-protein
CC interactions, drug-target interactions, and gene expression profiles
XX
SQ Sequence 863 AA;

Query Match 27.4%; Score 73.5; DB 6; Length 863;
Best Local Similarity 45.7%; Pred. No. 2.1;
Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

Qy 2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
::||:| || || :| | | : | |||||
Db 588 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 622

RESULT 10
ADC31071
ID ADC31071 standard; protein; 863 AA.
XX
AC ADC31071;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel polypeptide sequence, SEQ ID NO:1153.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;

```
KW   biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW   neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW   ulcers; osteoporosis; autoimmune disease; cancer;
KW   molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW   neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW   antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW   gene therapy; chromosome 22q13.1-13.2.
XX
OS   Homo sapiens.
XX
PN   WO2003029271-A2.
XX
PD   10-APR-2003.
XX
PF   24-SEP-2002; 2002WO-US030474.
XX
PR   24-SEP-2001; 2001US-0324631P.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang TY,  Zhang J,  Ren F,  Xue AJ,  Zhao QA,  Wang J,  Wehrman T;
PI   Zhou P,  Ghosh M,  Wang D,  Ma Y,  Asundi V,  Wang Z,  Weng G;
PI   Haley-Vicente D,  Drmanac RT;
XX
DR   WPI; 2003-371981/35.
DR   N-PSDB; ADC30100.
XX
PT   New polynucleotide and polypeptide useful for diagnosing, preventing or
PT   treating conditions such as neurodegenerative diseases, anemias, platelet
PT   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT   cancer.
XX
PS   Claim 20; SEQ ID NO 1153; 1185pp; English.
XX
CC   The invention relates to 971 novel human cDNA sequences (ADC29919-
CC   ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC   invention also relates to nucleic acid sequences over 99% identical with
CC   the novel human cDNAs. The invention additionally encompasses expression
CC   vectors and host cells comprising a nucleic acid of the invention; the
CC   recombinant production of a polypeptide of the invention; an antibody
CC   against a polypeptide of the invention; a method of detecting
CC   polynucleotides or polypeptides of the invention; and methods of
CC   identifying a compound which binds to a polypeptide of the invention. The
CC   invention further discloses methods of peventing, treating or
CC   ameliorating a medical condition; kits comprising polynucleotide probes
CC   and/or monoclonal antibodies for carrying out the methods of the
CC   invention; methods for the identification of compounds that modulate the
CC   expression or activity of the polynucleotide and/or polypeptide; and 767
CC   contig sequences corresponding to the cDNA sequences of the invention
CC   (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC   -ADC33394). The nucleic acids and polypeptides of the invention are
CC   useful in diagnostics, drug screening, forensics, gene mapping, in the
CC   identification of mutations responsible for genetic disorders or other
CC   traits, for assessing biodiversity, and in producing many other types of
CC   data and products dependent on DNA and amino acid sequences. They are
CC   also used for treating diseases such as Parkinson's disease, Alzheimer's
CC   disease and other neurodegenerative diseases, anaemia, platelet
CC   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC   cancer. The nucleic acids may also be used as hybridisation probes or
CC   primers, and in the recombinant production of a protein. The polypeptides
CC   are also useful in generating antibodies, as molecular weight markers,
CC   and as food supplements. The present sequence represents a specifically
CC   claimed human polypeptide sequence of the invention. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 863 AA;

  Query Match             27.4%;  Score 73.5;  DB 7;  Length 863;
  Best Local Similarity   45.7%;  Pred. No. 2.1;
  Matches   16;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db        588 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 622


RESULT 11
ADE48296
ID   ADE48296 standard; protein; 863 AA.
XX
AC   ADE48296;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human MICAL-1.
XX
KW   MICAL domain; calponin homology domain; LIM domain; plexin;
KW   Neuroprotective; Analgesic; Antiparkinsonian; Vulnerary; Tranquilizer;
KW   Vasotropic; Cardiant; Nootropic; Gene therapy; Antisense gene therapy;
KW   Cell therapy; Monooxygenase inhibitor; neurological condition.
XX
OS   Homo sapiens.
XX
PN   WO2003066821-A2.
XX
PD   14-AUG-2003.
XX
PF   04-FEB-2003; 2003WO-US003551.
XX
PR   04-FEB-2002; 2002US-0354178P.
PR   30-MAY-2002; 2002US-0384302P.
PR   13-JUN-2002; 2002US-0388325P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI   Kolodkin AL,  Terman JR,  Mao T,  Pasterkamp RJ,  Yu H;
XX
DR   WPI; 2003-731494/69.
DR   N-PSDB; ADE48295.
XX
PT   New polypeptide comprising an N-terminal MICAL domain, a calponin
PT   homology domain, a LIM domain, a proline rich region or a plexin
PT   interacting region, useful for treating a neurological condition, e.g.,
PT   Alzheimer's disease.
XX
PS   Claim 4; SEQ ID NO 2; 207pp; English.
XX
CC   The present sequence relates to an isolated polypeptide comprising an N-
CC   terminal MICAL domain, a calponin homology domain, a LIM domain, a
CC   proline rich region or a plexin interacting region, where the polypeptide
CC   has monooxygenase activity. The polypeptide is useful for treating a
CC   neurological condition, e.g., spinal cord injury, traumatic brain injury,
CC   neuropathic pain, Parkinson's disease, Amyotrophic Lateral Sclerosis,
CC   ischemic injury, Alzheimer's disease, multiple sclerosis or neuropathy
CC   resulting from a stroke. The present sequence represents a MICAL protein.
XX
SQ   Sequence 863 AA;

  Query Match             27.4%;  Score 73.5;  DB 7;  Length 863;
  Best Local Similarity   45.7%;  Pred. No. 2.1;
  Matches   16;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db        588 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 622


RESULT 12
ADE48308
ID   ADE48308 standard; protein; 863 AA.
XX
AC   ADE48308;
XX
DT   29-JAN-2004  (first entry)
```

XX
DE Human MICAL-1 like-protein.
XX
KW MICAL domain; calponin homology domain; LIM domain; plexin;
KW Neuroprotective; Analgesic; Antiparkinsonian; Vulnerary; Tranquilizer;
KW Vasotropic; Cardiant; Nootropic; Gene therapy; Antisense gene therapy;
KW Cell therapy; Monooxygenase inhibitor; neurological condition.
XX
OS Homo sapiens.
XX
PN WO2003066821-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003551.
XX
PR 04-FEB-2002; 2002US-0354178P.
PR 30-MAY-2002; 2002US-0384302P.
PR 13-JUN-2002; 2002US-0388325P.
XX
PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Kolodkin AL, Terman JR, Mao T, Pasterkamp RJ, Yu H;
XX
DR WPI; 2003-731494/69.
DR N-PSDB; ADE48307.
XX
PT New polypeptide comprising an N-terminal MICAL domain, a calponin
PT homology domain, a LIM domain, a proline rich region or a plexin
PT interacting region, useful for treating a neurological condition, e.g.,
PT Alzheimer's disease.
XX
PS Claim 72; SEQ ID NO 14; 207pp; English.
XX
CC The present sequence relates to an isolated polypeptide comprising an N-
CC terminal MICAL domain, a calponin homology domain, a LIM domain, a
CC proline rich region or a plexin interacting region, where the polypeptide
CC has monooxygenase activity. The polypeptide is useful for treating a
CC neurological condition, e.g., spinal cord injury, traumatic brain injury,
CC neuropathic pain, Parkinson's disease, Amyotrophic Lateral Sclerosis,
CC ischemic injury, Alzheimer's disease, multiple sclerosis or neuropathy
CC resulting from a stroke. The present sequence represents a MICAL protein.
XX
SQ Sequence 863 AA;

Query Match 27.4%; Score 73.5; DB 7; Length 863;
Best Local Similarity 45.7%; Pred. No. 2.1;
Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

Qy 2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
::||:| || || :| | | : | |||||
Db 588 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 622

RESULT 13
ABM82355
ID ABM82355 standard; protein; 863 AA.
XX
AC ABM82355;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO83279, SEQ:6048.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
DR N-PSDB; ACN40936.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 12; SEQ ID NO 6048; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 863 AA;

Query Match 27.4%; Score 73.5; DB 8; Length 863;
Best Local Similarity 45.7%; Pred. No. 2.1;
Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

Qy 2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
::||:| || || :| | | : | |||||
Db 588 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 622

RESULT 14
ABO76575
ID ABO76575 standard; protein; 173 AA.
XX
AC ABO76575;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #8750.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX

```
PF   18-FEB-1999;   99US-00252991.
XX
PR   18-FEB-1998;   98US-0074788P.
PR   27-JUL-1998;   98US-0094190P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Rubenfield MJ,  Nolling J,  Deloughery C,  Bush D;
XX
DR   WPI; 2003-615309/58.
DR   N-PSDB; ABD10146.
XX
PT   Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT   useful as molecular targets for diagnostics, prophylaxis and treatment of
PT   pathological conditions resulting from bacterial infection.
XX
PS   Disclosure; SEQ ID NO 25321; 455pp; English.
XX
CC   The invention relates to Pseudomonas aeruginosa polypeptides and the
CC   polynucleotides encoding them. The sequences are useful in diagnosis and
CC   therapy of pathological conditions, as molecular targets for diagnostics,
CC   prophylaxis and treatment of pathological conditions resulting from a
CC   bacterial infection, for evaluating a compound, such as a polypeptide,
CC   for the ability to bind a P. aeruginosa nucleic acid, as components of
CC   effective antibacterial targets, as targets for antibacterial drugs,
CC   including anti-P. aeruginosa drugs, as templates for recombinant
CC   production of P. aeruginosa-derived peptides or polypeptides, as target
CC   components for diagnosis and/or treatment of P. aeruginosa-caused
CC   infection, and in detection of P. aeruginosa sequences or other sequences
CC   of Pseudomonas species using biochip technology. Sequences ABO67826-
CC   ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format from USPTO at
CC   seqdata.uspto.gov/sequence.html
XX
SQ   Sequence 173 AA;

  Query Match             27.1%;  Score 72.5;  DB 7;  Length 173;
  Best Local Similarity   40.5%;  Pred. No. 0.41;
  Matches   17;  Conservative    4;  Mismatches   18;  Indels    3;  Gaps    1;

Qy          5 PRPGAPCRAGAPTIVLTSGRRQTLSHG---SSSPARATLGKP 43
              |||   ||||| |    : ||:    |   : :| ||  | |
Db         44 PRPARACRAGAGTDATGAARRRFAGAGPRRAGTPTRAPAGSP 85


RESULT 15
AAG28488
ID   AAG28488 standard; protein; 76 AA.
XX
AC   AAG28488;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Zea mays protein fragment SEQ ID NO: 33724.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence; corn.
XX
OS   Zea mays subsp. mays.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
```

PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 26.7%; Score 71.5; DB 3; Length 76;
Best Local Similarity 34.0%; Pred. No. 0.21;
Matches 17; Conservative 10; Mismatches 20; Indels 3; Gaps 1;

```
Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSL 51
              |:  || :| || : | : ||  : | :|| : |  :   |   || |::
Db         15 ASSSRPSSPARARSXTSLTTSLTQSTTTHGPTPPTSSPATK---LNGYAM 61
```

Search completed: April 27, 2006, 15:14:40
Job time : 223.235 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 27, 2006, 15:20:00 ; Search time 50.4706 Seconds
(without alignments)
85.181 Million cell updates/sec

Title: US-09-819-144A-2
Perfect score: 268
Sequence: 1 MAAGPRPGAPCRAGAPTIVL.....SSPARATLGKPLVLNDYSLN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	268	100.0	52	2	US-09-042-225-2	Sequence 2, Appli
2	74	27.6	16	2	US-09-042-225-8	Sequence 8, Appli
3	72.5	27.1	173	2	US-09-252-991A-25321	Sequence 25321, A
4	67.5	25.2	137	2	US-09-252-991A-30540	Sequence 30540, A
5	66	24.6	228	2	US-09-252-991A-28381	Sequence 28381, A
6	65	24.3	166	2	US-09-252-991A-16776	Sequence 16776, A
7	65	24.3	205	2	US-09-252-991A-31167	Sequence 31167, A
8	65	24.3	345	2	US-09-252-991A-25768	Sequence 25768, A
9	64.5	24.1	146	2	US-09-252-991A-25930	Sequence 25930, A
10	64.5	24.1	149	2	US-09-252-991A-30866	Sequence 30866, A
11	63.5	23.7	721	2	US-09-252-991A-28293	Sequence 28293, A
12	63	23.5	128	2	US-10-104-047-3112	Sequence 3112, Ap
13	63	23.5	613	2	US-09-252-991A-19955	Sequence 19955, A
14	63	23.5	1129	2	US-09-252-991A-22330	Sequence 22330, A
15	62	23.1	256	2	US-09-252-991A-25404	Sequence 25404, A
16	62	23.1	328	2	US-09-252-991A-19582	Sequence 19582, A
17	62	23.1	863	2	US-09-252-991A-26099	Sequence 26099, A
18	61.5	22.9	146	2	US-09-949-016-8300	Sequence 8300, Ap
19	61.5	22.9	367	2	US-09-252-991A-19910	Sequence 19910, A
20	61.5	22.9	574	2	US-09-252-991A-30868	Sequence 30868, A
21	61	22.8	170	2	US-09-252-991A-20706	Sequence 20706, A
22	61	22.8	208	2	US-09-252-991A-17849	Sequence 17849, A
23	61	22.8	679	2	US-09-252-991A-27111	Sequence 27111, A
24	60.5	22.6	155	2	US-09-252-991A-32893	Sequence 32893, A
25	60.5	22.6	240	2	US-09-252-991A-28112	Sequence 28112, A
26	60	22.4	343	2	US-09-252-991A-27631	Sequence 27631, A
27	60	22.4	755	2	US-09-902-540-11169	Sequence 11169, A
28	59.5	22.2	158	2	US-09-252-991A-24896	Sequence 24896, A
29	59.5	22.2	293	2	US-09-252-991A-32060	Sequence 32060, A
30	59.5	22.2	395	2	US-09-270-767-44351	Sequence 44351, A
31	59.5	22.2	413	2	US-09-604-231-32	Sequence 32, Appl
32	59.5	22.2	536	2	US-09-359-167-8	Sequence 8, Appli
33	59.5	22.2	606	2	US-09-486-382B-11	Sequence 11, Appl
34	59.5	22.2	683	2	US-09-604-231-30	Sequence 30, Appl
35	59.5	22.2	1031	2	US-09-252-991A-18365	Sequence 18365, A
36	59	22.0	162	2	US-09-252-991A-32101	Sequence 32101, A
37	59	22.0	439	2	US-09-252-991A-17127	Sequence 17127, A
38	59	22.0	770	2	US-09-784-316-5	Sequence 5, Appli
39	59	22.0	770	2	US-10-229-124-5	Sequence 5, Appli
40	58.5	21.8	141	2	US-09-252-991A-23685	Sequence 23685, A
41	58.5	21.8	377	2	US-09-252-991A-32966	Sequence 32966, A
42	58.5	21.8	515	2	US-09-252-991A-23632	Sequence 23632, A
43	58.5	21.8	802	2	US-09-252-991A-25050	Sequence 25050, A
44	58.5	21.8	929	2	US-09-252-991A-19203	Sequence 19203, A
45	58	21.6	117	2	US-09-489-847-199	Sequence 199, App

ALIGNMENTS

RESULT 1
US-09-042-225-2
; Sequence 2, Application US/09042225A
; Patent No. 6207812
; GENERAL INFORMATION:
; APPLICANT: Terek, Richard M.
; TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES
; FILE REFERENCE: 04930/021001
; CURRENT APPLICATION NUMBER: US/09/042,225A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-042-225-2

Query Match 100.0%; Score 268; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 5.9e-27;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSLN 52
||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MAAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSLN 52

RESULT 2
US-09-042-225-8
; Sequence 8, Application US/09042225A
; Patent No. 6207812
; GENERAL INFORMATION:
; APPLICANT: Terek, Richard M.
; TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES
; FILE REFERENCE: 04930/021001
; CURRENT APPLICATION NUMBER: US/09/042,225A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-042-225-8

Query Match 27.6%; Score 74; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 RRQTLSHGSSSPARA 38

```
                ||||||||||||||||
Db          1 RRQTLSHGSSSPARA 15


RESULT 3
US-09-252-991A-25321
; Sequence 25321, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 25321
;   LENGTH: 173
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25321

  Query Match             27.1%;  Score 72.5;  DB 2;  Length 173;
  Best Local Similarity   40.5%;  Pred. No. 0.17;
  Matches   17;  Conservative    4;  Mismatches   18;  Indels    3;  Gaps    1;

Qy          5 PRPGAPCRAGAPTIVLTSGRRQTLSHG---SSSPARATLGKP 43
              |||   ||||| |    : ||:    |   : :| ||  | |
Db         44 PRPARACRAGAGTDATGAARRRFAGAGPRRAGTPTRAPAGSP 85


RESULT 4
US-09-252-991A-30540
; Sequence 30540, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 30540
;   LENGTH: 137
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30540

  Query Match             25.2%;  Score 67.5;  DB 2;  Length 137;
  Best Local Similarity   36.0%;  Pred. No. 0.54;
  Matches   18;  Conservative    5;  Mismatches   12;  Indels   15;  Gaps    2;

Qy          2 AAGPRPG-------------APCRAGAPTIVLTSGRRQTLSHGSSSPAR 37
              |  ||||               || || |:     | |:::: || ||||
Db         44 APAPRPGRRGPSIDRRGSPGPRPCLAG-PSAASPCGPRRSIAGGSRSPAR 92


RESULT 5
US-09-252-991A-28381
; Sequence 28381, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 28381
;   LENGTH: 228
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28381

  Query Match             24.6%;  Score 66;  DB 2;  Length 228;
  Best Local Similarity   38.1%;  Pred. No. 1.5;
  Matches   16;  Conservative    3;  Mismatches    9;  Indels   14;  Gaps    2;

Qy          3 AGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATL-GKP 43
              ||  ||||||: :|             || | |    | |:|
Db         77 AGKAPGAPCRSASP------------HGRSCPPPGRLRGRP 105


RESULT 6
US-09-252-991A-16776
; Sequence 16776, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 16776
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16776

  Query Match             24.3%;  Score 65;  DB 2;  Length 166;
  Best Local Similarity   46.2%;  Pred. No. 1.4;
  Matches   18;  Conservative    4;  Mismatches    5;  Indels   12;  Gaps    2;

Qy          4 GPRPGAPCRAGAPTIVLTSGRRQTL----SHGSSSPARA 38
              |||||:| |         ::||| :     | || |||||
Db        115 GPRPGSPDR--------SAGRRGSSPPPGSGGSPSPARA 145


RESULT 7
US-09-252-991A-31167
; Sequence 31167, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 31167
;   LENGTH: 205
```

```
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31167

  Query Match             24.3%;  Score 65;  DB 2;  Length 205;
  Best Local Similarity   32.1%;  Pred. No. 1.8;
  Matches   18;  Conservative    6;  Mismatches   18;  Indels   14;  Gaps    2;

Qy          2 AAGPRPGAPCRAGAPTIV---------LTSGRRQTLSHGSSSP-----ARATLGKP 43
              | || |  | | | |  |           :|||:  : ||::|     |    |:|
Db        142 APGPGPARPVRRGGPPAVRRGECFPVRRRAGRRRDRASGSAAPPGGDRAAGRAGRP 197


RESULT 8
US-09-252-991A-25768
; Sequence 25768, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 25768
;   LENGTH: 345
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25768

  Query Match             24.3%;  Score 65;  DB 2;  Length 345;
  Best Local Similarity   40.5%;  Pred. No. 3.4;
  Matches   15;  Conservative    5;  Mismatches   13;  Indels    4;  Gaps    1;

Qy          5 PRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLG 41
              | |||| | |    :    ||:  :|   ||:| |:|
Db        150 PTPGAPTRIG----LFCRRRRRMSTHWRISPSRPTMG 182


RESULT 9
US-09-252-991A-25930
; Sequence 25930, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 25930
;   LENGTH: 146
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25930

  Query Match             24.1%;  Score 64.5;  DB 2;  Length 146;
  Best Local Similarity   35.4%;  Pred. No. 1.4;
  Matches   17;  Conservative    5;  Mismatches   19;  Indels    7;  Gaps    1;

Qy          3 AGPRPGAPCRAGAPTIVLTSGRRQTLSHGS-------SSPARATLGKP 43
              | || |:|  ||  :| : || | | :  |       |||    :  |
Db          8 ASPRSGSPMTAGRWSIPMGSGTRSTATSSSPPAGYSRSSPPSTAVASP 55


RESULT 10
US-09-252-991A-30866
; Sequence 30866, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 30866
;   LENGTH: 149
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30866

  Query Match             24.1%;  Score 64.5;  DB 2;  Length 149;
  Best Local Similarity   41.7%;  Pred. No. 1.4;
  Matches   15;  Conservative    2;  Mismatches   10;  Indels    9;  Gaps    1;

Qy          4 GPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARAT 39
              | |||  ||         :||||   |   || ||:
Db         99 GARPGGSCR---------NGRRQGSGHAWGSPQRAS 125


RESULT 11
US-09-252-991A-28293
; Sequence 28293, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 28293
;   LENGTH: 721
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28293

  Query Match             23.7%;  Score 63.5;  DB 2;  Length 721;
  Best Local Similarity   36.8%;  Pred. No. 13;
  Matches   21;  Conservative    4;  Mismatches   17;  Indels   15;  Gaps    2;

Qy          4 GPRP-----GAPCRAGAPTIVLTSGRRQTLSHG----------SSSPARATLGKPLV 45
              | ||     |||| | | |    |||| |   |          :::  | :|| | |
Db        242 GRRPDPRQAGAPCTARATTGRAGSGRRLTSRPGCRRWRTCGRPANAACRRSLGSPSV 298


RESULT 12
US-10-104-047-3112
; Sequence 3112, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. 6943241el full length cDNA
;  FILE REFERENCE: H1-A0105
;  CURRENT APPLICATION NUMBER: US/10/104,047
```

```
;  CURRENT FILING DATE:  2002-03-25
;  PRIOR APPLICATION NUMBER:
;  PRIOR FILING DATE:
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3112
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-104-047-3112

  Query Match             23.5%;  Score 63;  DB 2;  Length 128;
  Best Local Similarity   35.6%;  Pred. No. 1.9;
  Matches   16;  Conservative    4;  Mismatches   23;  Indels    2;  Gaps    1;

Qy          4 GPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLND 48
              ||  |     | |    |:|    |  |:| | :  || || | :
Db          2 GPAAGHAAPRGPPVPATTAG--AALRAGASEPRQTQLGAPLALGE 44


RESULT 13
US-09-252-991A-19955
; Sequence 19955, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 19955
;   LENGTH: 613
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19955

  Query Match             23.5%;  Score 63;  DB 2;  Length 613;
  Best Local Similarity   41.0%;  Pred. No. 12;
  Matches   16;  Conservative    5;  Mismatches   16;  Indels    2;  Gaps    1;

Qy          5 PRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKP 43
              |||||  |  |      :|| :  :|| |:   | ||:|
Db        165 PRPGADPRGPARQAARIAGRHR--AHGPSANPDAPLGRP 201


RESULT 14
US-09-252-991A-22330
; Sequence 22330, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 22330
;   LENGTH: 1129
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22330

  Query Match             23.5%;  Score 63;  DB 2;  Length 1129;
  Best Local Similarity   44.8%;  Pred. No. 25;
  Matches   13;  Conservative    6;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         13 AGAPTIVLTSGRRQTLSHGSSSPARATLG 41
              :|: |:|| || |  | ||:::||    |
Db        938 SGSGTLVLASGARIDLRHGTAAPAEQVDG 966


RESULT 15
US-09-252-991A-25404
; Sequence 25404, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 25404
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25404

  Query Match             23.1%;  Score 62;  DB 2;  Length 256;
  Best Local Similarity   33.3%;  Pred. No. 5.7;
  Matches   20;  Conservative    6;  Mismatches   14;  Indels   20;  Gaps    3;

Qy          2 AAGPRPGA-------------PC------RAGAPTIVLTSGRRQTLSHGSSSPARATLGK 42
              | | ||||             ||      | ||| :  |:|||:  :  :  | ||  |:
Db         84 ANGCRPGAAALRPVRGAPGRWPCRIAGRGRPGAP-VAGTAGRRRATASRARPPPRADRGR 142


Search completed: April 27, 2006, 15:21:43
Job time : 51.4706 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 27, 2006, 15:35:30 ; Search time 182 Seconds
(without alignments)
119.380 Million cell updates/sec

Title: US-09-819-144A-2
Perfect score: 268
Sequence: 1 MAAGPRPGAPCRAGAPTIVL.....SSPARATLGKPLVLNDYSLN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	268	100.0	52	3	US-09-819-144A-2	Sequence 2, Appli
2	79	29.5	229	4	US-10-425-115-222906	Sequence 222906,
3	77	28.7	190	4	US-10-437-963-198300	Sequence 198300,
4	74	27.6	16	3	US-09-819-144A-8	Sequence 8, Appli
5	73.5	27.4	791	4	US-10-170-385-57	Sequence 57, Appl
6	73.5	27.4	863	4	US-10-359-012-2	Sequence 2, Appli
7	73.5	27.4	863	4	US-10-359-012-14	Sequence 14, Appl
8	71.5	26.7	240	4	US-10-425-115-226148	Sequence 226148,
9	71	26.5	19608	4	US-10-084-846A-8	Sequence 8, Appli
10	68	25.4	315	4	US-10-080-170-423	Sequence 423, App
11	68	25.4	315	4	US-10-080-170-423	Sequence 423, App
12	68	25.4	315	4	US-10-468-356-423	Sequence 423, App
13	67	25.0	219	4	US-10-425-115-262578	Sequence 262578,
14	67	25.0	1321	5	US-10-840-512-117	Sequence 117, App
15	66.5	24.8	185	4	US-10-425-115-341679	Sequence 341679,
16	65.5	24.4	923	5	US-10-756-149-5246	Sequence 5246, Ap
17	64.5	24.1	137	4	US-10-767-701-33507	Sequence 33507, A
18	64.5	24.1	309	5	US-10-450-763-43334	Sequence 43334, A
19	64.5	24.1	596	4	US-10-156-761-8931	Sequence 8931, Ap
20	64	23.9	165	4	US-10-425-115-270067	Sequence 270067,
21	64	23.9	359	4	US-10-437-963-190864	Sequence 190864,
22	63.5	23.7	74	4	US-10-425-115-263976	Sequence 263976,
23	63.5	23.7	87	4	US-10-425-115-313357	Sequence 313357,
24	63.5	23.7	223	4	US-10-425-114-62626	Sequence 62626, A
25	63	23.5	128	4	US-10-104-047-3112	Sequence 3112, Ap
26	63	23.5	128	4	US-10-108-260A-2594	Sequence 2594, Ap
27	62.5	23.3	118	4	US-10-424-599-271704	Sequence 271704,
28	62.5	23.3	197	4	US-10-437-963-128246	Sequence 128246,
29	62.5	23.3	208	4	US-10-437-963-176133	Sequence 176133,
30	62.5	23.3	524	5	US-10-450-763-38710	Sequence 38710, A
31	62	23.1	79	4	US-10-425-115-259836	Sequence 259836,
32	62	23.1	175	4	US-10-425-115-210919	Sequence 210919,
33	62	23.1	177	4	US-10-437-963-109208	Sequence 109208,
34	62	23.1	259	4	US-10-443-622-59	Sequence 59, Appl
35	62	23.1	259	5	US-10-974-440-30	Sequence 30, Appl
36	61.5	22.9	135	4	US-10-425-115-350149	Sequence 350149,
37	61.5	22.9	154	4	US-10-767-701-55508	Sequence 55508, A
38	61.5	22.9	170	3	US-09-864-408A-6684	Sequence 6684, Ap
39	61.5	22.9	240	4	US-10-425-115-326091	Sequence 326091,
40	61	22.8	79	4	US-10-425-115-357643	Sequence 357643,
41	61	22.8	169	4	US-10-767-701-42973	Sequence 42973, A
42	61	22.8	176	4	US-10-425-115-203617	Sequence 203617,
43	61	22.8	241	4	US-10-425-115-245202	Sequence 245202,
44	61	22.8	595	4	US-10-156-761-10072	Sequence 10072, A
45	61	22.8	656	4	US-10-108-260A-2611	Sequence 2611, Ap

ALIGNMENTS

RESULT 1
US-09-819-144A-2
; Sequence 2, Application US/09819144A
; Publication No. US20010016649A1
; GENERAL INFORMATION:
; APPLICANT: Terek, Richard M.
; TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES
; FILE REFERENCE: 21486-021DIV
; CURRENT APPLICATION NUMBER: US/09/819,144A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-819-144A-2

Query Match 100.0%; Score 268; DB 3; Length 52;
Best Local Similarity 100.0%; Pred. No. 2.6e-24;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSLN 52
||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MAAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSLN 52

RESULT 2
US-10-425-115-222906
; Sequence 222906, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 222906
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(229)
; OTHER INFORMATION: unsure at all Xaa locations

```
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_134881C.1.pep
US-10-425-115-222906

  Query Match             29.5%;  Score 79;  DB 4;  Length 229;
  Best Local Similarity   43.8%;  Pred. No. 0.38;
  Matches   21;  Conservative    5;  Mismatches   16;  Indels    6;  Gaps    2;

Qy          4 GPRPGAPCRAG--APTIVLTSGRRQTLSHG----SSSPARATLGKPLV 45
              |  |   || |  ||| :|   ||:|| ||    ||||   |: : |:
Db         32 GQLPAKVCRTGGSAPTRILXRRRRRTLHHGSPELSSSPPPCTICRRLL 79


RESULT 3
US-10-437-963-198300
; Sequence 198300, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198300
;   LENGTH: 190
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_93974C.1.pep
US-10-437-963-198300

  Query Match             28.7%;  Score 77;  DB 4;  Length 190;
  Best Local Similarity   35.1%;  Pred. No. 0.54;
  Matches   20;  Conservative   11;  Mismatches   16;  Indels   10;  Gaps    3;

Qy          1 MAAGPRPGAP------CRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSL 51
              :|  | ||||      |  | ||:::| |    :|:|::| : | |   |::   ||
Db        131 LAPAP-PGAPSPGITICSGGPPTVIMTPG---VISYGAASRSSANLSSSLLVAMVSL 183


RESULT 4
US-09-819-144A-8
; Sequence 8, Application US/09819144A
; Publication No. US20010016649A1
; GENERAL INFORMATION:
;  APPLICANT: Terek, Richard  M.
;  TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES
;  FILE REFERENCE: 21486-021DIV
;  CURRENT APPLICATION NUMBER: US/09/819,144A
;  CURRENT FILING DATE:  2001-03-27
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-819-144A-8

  Query Match             27.6%;  Score 74;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.078;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         24 RRQTLSHGSSSPARA 38
              |||||||||||||||
Db          1 RRQTLSHGSSSPARA 15


RESULT 5
US-10-170-385-57
; Sequence 57, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
;  APPLICANT: Ward, Neil Raymond
;  APPLICANT:  Mundy, Christopher Robert
;  APPLICANT:  Kan, On
;  APPLICANT:  Harris, Robert Alan
;  APPLICANT:  White, Jonathan
;  APPLICANT:  Binley, Katie Mary
;  APPLICANT:  Rayner, William Nigel
;  APPLICANT:  Naylor, Stuart
;  APPLICANT:  Kingsman, Susan Mary
;  APPLICANT:  Krige, David
;  TITLE OF INVENTION: ANALYSIS METHOD
;  FILE REFERENCE: 532682000100
;  CURRENT APPLICATION NUMBER: US/10/170,385
;  CURRENT FILING DATE:  2002-06-12
;  PRIOR APPLICATION NUMBER: PCT/GB02/01662
;  PRIOR FILING DATE: 2002-04-08
;  PRIOR APPLICATION NUMBER: PCT/GB01/05458
;  PRIOR FILING DATE: 2001-12-10
;  NUMBER OF SEQ ID NOS: 549
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 57
;   LENGTH: 791
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-170-385-57

  Query Match             27.4%;  Score 73.5;  DB 4;  Length 791;
  Best Local Similarity   45.7%;  Pred. No. 6.8;
  Matches   16;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db        516 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 550


RESULT 6
US-10-359-012-2
; Sequence 2, Application US/10359012
; Publication No. US20030232419A1
; GENERAL INFORMATION:
;  APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
;  APPLICANT:  KOLODKIN, Alex L.
;  APPLICANT:  TERMAN, Jon R.
;  APPLICANT:  MAO, Tianyi
;  APPLICANT:  PASTERKAMP, Ronald J.
;  APPLICANT:  YU, Hung-Hsiang
;  TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPT
;  TITLE OF INVENTION:  AND METHODS OF USING THE SAME
;  FILE REFERENCE: JHU1840-3
;  CURRENT APPLICATION NUMBER: US/10/359,012
;  CURRENT FILING DATE:  2003-02-04
;  PRIOR APPLICATION NUMBER: US 60/388,325
;  PRIOR FILING DATE: 2002-06-13
;  PRIOR APPLICATION NUMBER: US 60/384,302
;  PRIOR FILING DATE: 2002-05-30
;  PRIOR APPLICATION NUMBER: US 60/354,178
;  PRIOR FILING DATE: 2002-02-04
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 863
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-359-012-2
```

```
  Query Match             27.4%;  Score 73.5;  DB 4;  Length 863;
  Best Local Similarity   45.7%;  Pred. No. 7.5;
  Matches   16;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db        588 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 622


RESULT 7
US-10-359-012-14
; Sequence 14, Application US/10359012
; Publication No. US20030232419A1
; GENERAL INFORMATION:
;  APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
;  APPLICANT:  KOLODKIN, Alex L.
;  APPLICANT:  TERMAN, Jon R.
;  APPLICANT:  MAO, Tianyi
;  APPLICANT:  PASTERKAMP, Ronald J.
;  APPLICANT:  YU, Hung-Hsiang
;  TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPT
;  TITLE OF INVENTION:  AND METHODS OF USING THE SAME
;  FILE REFERENCE: JHU1840-3
;  CURRENT APPLICATION NUMBER: US/10/359,012
;  CURRENT FILING DATE:  2003-02-04
;  PRIOR APPLICATION NUMBER: US 60/388,325
;  PRIOR FILING DATE: 2002-06-13
;  PRIOR APPLICATION NUMBER: US 60/384,302
;  PRIOR FILING DATE: 2002-05-30
;  PRIOR APPLICATION NUMBER: US 60/354,178
;  PRIOR FILING DATE: 2002-02-04
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 863
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-359-012-14

  Query Match             27.4%;  Score 73.5;  DB 4;  Length 863;
  Best Local Similarity   45.7%;  Pred. No. 7.5;
  Matches   16;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db        588 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 622


RESULT 8
US-10-425-115-226148
; Sequence 226148, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 226148
;   LENGTH: 240
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_137836C.1.pep
US-10-425-115-226148

  Query Match             26.7%;  Score 71.5;  DB 4;  Length 240;
  Best Local Similarity   35.1%;  Pred. No. 3.1;
  Matches   26;  Conservative    4;  Mismatches   19;  Indels   25;  Gaps    3;

Qy          1 MAAGPRPGA--PCRAGA----------------------PTIVLTSGRRQTLSHGSSSPA 36
                :|  |||||  | |  |                      |:  | | ||:  | |  ||
Db         64 LAGPPRPGARPPTRFAADPRVDSCSTVAHFGSCPAPSPHPSQPLLS-RRRLPSGGPRSPR 122

Qy         37 RATLGKPLVLNDYS 50
                || ||||  :| |
Db        123 TATGGKPLAAHDIS 136


RESULT 9
US-10-084-846A-8
; Sequence 8, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
;  APPLICANT: WEITNAUER, GABRIELE
;  APPLICANT:  MUHLENWEG, AGNES
;  APPLICANT:  TREFZER, AXEL
;  APPLICANT:  BECHTHOLD, ANDREAS
;  TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
;  FILE REFERENCE: 1974-005
;  CURRENT APPLICATION NUMBER: US/10/084,846A
;  CURRENT FILING DATE:  2003-02-25
;  PRIOR APPLICATION NUMBER: PCT/EP01/09815
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: DE 101 09 166.4
;  PRIOR FILING DATE: 2001-02-25
;  NUMBER OF SEQ ID NOS: 120
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 8
;   LENGTH: 19608
;   TYPE: PRT
;   ORGANISM: Streptomyces viridochromogenes
;   FEATURE:
;   OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
;   OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.
US-10-084-846A-8

  Query Match             26.5%;  Score 71;  DB 4;  Length 19608;
  Best Local Similarity   39.6%;  Pred. No. 4.8e+02;
  Matches   19;  Conservative    2;  Mismatches   15;  Indels   12;  Gaps    2;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTLSHGS------SSPARATLGKP 43
                |  ||||  || |||      |:|   | |:      ||| |     |
Db       3641 ARPPRPGRHCRGGAP------GKRSGRSRGTAAALADSSPGRTARSGP 3682


RESULT 10
US-10-080-170-423
; Sequence 423, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
;  APPLICANT: COLE, S.T.
;  TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
;  TITLE OF INVENTION:  IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
;  TITLE OF INVENTION:  TREATMENT OF MYCOBACTERIOSES
;  FILE REFERENCE: 03495.0218
;  CURRENT APPLICATION NUMBER: US/10/080,170
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: 60/270,123
;  PRIOR FILING DATE: 2001-02-22
;  NUMBER OF SEQ ID NOS: 652
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 423
;   LENGTH: 315
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis
US-10-080-170-423

  Query Match             25.4%;  Score 68;  DB 4;  Length 315;
  Best Local Similarity   28.2%;  Pred. No. 11;
  Matches   20;  Conservative    5;  Mismatches   14;  Indels   32;  Gaps    2;
```

```
Qy          1 MAAGPRPGAPC-------------------------RAGAPT------IVLTSGRRQTL 28
              : || ||| |                         | |  |      :  |:|| |
Db         51 LVAGSRPGGPIIGYLNLSPPRGAGGAMAELVVHPQSRRRGIGTAMARAALAKTAGRNQFW 110

Qy         29 SHGSSSPARAT 39
              :||:  |||||
Db        111 AHGTLDPARAT 121
```

RESULT 11
US-10-080-170-423
; Sequence 423, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 423
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-423

Query Match 25.4%; Score 68; DB 4; Length 315;
Best Local Similarity 28.2%; Pred. No. 11;
Matches 20; Conservative 5; Mismatches 14; Indels 32; Gaps 2;

```
Qy          1 MAAGPRPGAPC-------------------------RAGAPT------IVLTSGRRQTL 28
              : || ||| |                         | |  |      :  |:|| |
Db         51 LVAGSRPGGPIIGYLNLSPPRGAGGAMAELVVHPQSRRRGIGTAMARAALAKTAGRNQFW 110

Qy         29 SHGSSSPARAT 39
              :||:  |||||
Db        111 AHGTLDPARAT 121
```

RESULT 12
US-10-468-356-423
; Sequence 423, Application US/10468356
; Publication No. US20040197896A1
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENEOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 05394.0019
; CURRENT APPLICATION NUMBER: US/10/468,356
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: 10/080,170
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 655
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 423
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-468-356-423

Query Match 25.4%; Score 68; DB 4; Length 315;
Best Local Similarity 28.2%; Pred. No. 11;
Matches 20; Conservative 5; Mismatches 14; Indels 32; Gaps 2;

```
Qy          1 MAAGPRPGAPC-------------------------RAGAPT------IVLTSGRRQTL 28
              : || ||| |                         | |  |      :  |:|| |
Db         51 LVAGSRPGGPIIGYLNLSPPRGAGGAMAELVVHPQSRRRGIGTAMARAALAKTAGRNQFW 110

Qy         29 SHGSSSPARAT 39
              :||:  |||||
Db        111 AHGTLDPARAT 121
```

RESULT 13
US-10-425-115-262578
; Sequence 262578, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 262578
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_171081C.1.pep
US-10-425-115-262578

Query Match 25.0%; Score 67; DB 4; Length 219;
Best Local Similarity 38.1%; Pred. No. 9.7;
Matches 24; Conservative 3; Mismatches 20; Indels 16; Gaps 4;

```
Qy          3 AGPRPGAP-CRAGAPTIVLTSGRRQTLSHGS-------SSPA-------RATLGKPLVLN 47
              ||  |||| || |       | |  :| || :          ||        || :| |||
Db        152 AGLAPGAPSCRKGGCRAARTGGPPKT-SHAARAAFVFWGGPAFAPVRRLRAEVGAPLVRE 210

Qy         48 DYS 50
              | |
Db        211 DLS 213
```

RESULT 14
US-10-840-512-117
; Sequence 117, Application US/10840512
; Publication No. US20050125852A1
; GENERAL INFORMATION:
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: MANNING, GERARD
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: GRIGORIEV, IGOR
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-1455
; CURRENT APPLICATION NUMBER: US/10/840,512
; CURRENT FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: 60/469,014
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 239
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 117
; LENGTH: 1321
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-840-512-117

Query Match 25.0%; Score 67; DB 5; Length 1321;
Best Local Similarity 39.5%; Pred. No. 72;
Matches 17; Conservative 8; Mismatches 10; Indels 8; Gaps 2;

```
Qy          7 PGAPCRAGAPTIVL-----TSGRRQTLSHGSSSPARATLGKPL 44
```

```
                ||:||   :|:: |      || |::|  |: ||   || :||
Db         29 PGSPCSPCSPSLGLQPWSCRSGNRKSLVVGTPSP---TLSRPL 68
```

RESULT 15
US-10-425-115-341679
; Sequence 341679, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 341679
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(185)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_74777C.1.pep
US-10-425-115-341679

Query Match 24.8%; Score 66.5; DB 4; Length 185;
Best Local Similarity 42.0%; Pred. No. 9.2;
Matches 21; Conservative 3; Mismatches 19; Indels 7; Gaps 2;

```
Qy          1 MAAGPRPGAPCRAG----APTIVLTS---GRRQTLSHGSSSPARATLGKP 43
              :| |   || |||     ||   |:|   |||   : ||||   ||   |
Db         61 IAPGTTRGAACRASSWTTAPASGLSSRSRGRRSPCAGGSSSTTAATSAPP 110
```

Search completed: April 27, 2006, 15:40:41
Job time : 183 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 27, 2006, 15:36:55 ; Search time 26 Seconds
(without alignments)
90.956 Million cell updates/sec

Title: US-09-819-144A-2
Perfect score: 268
Sequence: 1 MAAGPRPGAPCRAGAPTIVL.....SSPARATLGKPLVLNDYSLN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 232119

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /SIDS5/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /SIDS5/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /SIDS5/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /SIDS5/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /SIDS5/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /SIDS5/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /SIDS5/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /SIDS5/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	73.5	27.4	863	7	US-11-169-041-167	Sequence 167, App
2	70.5	26.3	219	7	US-11-096-568A-18567	Sequence 18567, A
3	70	26.1	497	7	US-11-096-568A-3275	Sequence 3275, Ap
4	70	26.1	498	7	US-11-096-568A-3274	Sequence 3274, Ap
5	70	26.1	498	7	US-11-096-568A-3276	Sequence 3276, Ap
6	70	26.1	550	7	US-11-096-568A-3273	Sequence 3273, Ap
7	63	23.5	128	7	US-11-072-512-3112	Sequence 3112, Ap
8	61.5	22.9	251	7	US-11-096-568A-12554	Sequence 12554, A
9	61.5	22.9	406	7	US-11-096-568A-12552	Sequence 12552, A
10	60.5	22.6	950	6	US-10-501-035-357	Sequence 357, App
11	58.5	21.8	265	7	US-11-051-724-38	Sequence 38, Appl
12	58.5	21.8	375	7	US-11-096-568A-23618	Sequence 23618, A
13	58	21.6	117	7	US-11-229-769-199	Sequence 199, App
14	58	21.6	127	7	US-11-229-769-351	Sequence 351, App
15	58	21.6	225	7	US-11-229-769-238	Sequence 238, App
16	58	21.6	225	7	US-11-229-769-353	Sequence 353, App
17	57.5	21.5	108	7	US-11-096-568A-17648	Sequence 17648, A
18	57.5	21.5	642	6	US-10-491-468-28	Sequence 28, Appl
19	57.5	21.5	1614	7	US-11-108-528-82	Sequence 82, Appl
20	57	21.3	255	7	US-11-096-568A-13556	Sequence 13556, A
21	57	21.3	260	7	US-11-226-657-65	Sequence 65, Appl
22	57	21.3	295	7	US-11-096-568A-13555	Sequence 13555, A
23	57	21.3	311	6	US-10-330-773-364	Sequence 364, App
24	57	21.3	575	6	US-10-131-826A-128	Sequence 128, App
25	57	21.3	575	6	US-10-973-115B-128	Sequence 128, App
26	57	21.3	575	6	US-10-137-873A-128	Sequence 128, App
27	57	21.3	575	6	US-10-152-370-128	Sequence 128, App
28	57	21.3	575	7	US-11-290-153-128	Sequence 128, App
29	56.5	21.1	125	7	US-11-096-568A-26408	Sequence 26408, A
30	56.5	21.1	186	7	US-11-096-568A-26407	Sequence 26407, A
31	56.5	21.1	212	7	US-11-096-568A-26406	Sequence 26406, A
32	56.5	21.1	512	7	US-11-087-099-6798	Sequence 6798, Ap
33	56	20.9	142	7	US-11-096-568A-21109	Sequence 21109, A
34	56	20.9	184	7	US-11-096-568A-11705	Sequence 11705, A
35	56	20.9	222	7	US-11-096-568A-19362	Sequence 19362, A
36	56	20.9	251	7	US-11-098-686-10141	Sequence 10141, A
37	56	20.9	811	7	US-11-072-512-3559	Sequence 3559, Ap
38	56	20.9	1306	6	US-10-995-561-1027	Sequence 1027, Ap
39	55	20.5	221	7	US-11-096-568A-17196	Sequence 17196, A
40	55	20.5	233	7	US-11-096-568A-17195	Sequence 17195, A
41	55	20.5	296	7	US-11-096-568A-34127	Sequence 34127, A
42	55	20.5	323	7	US-11-096-568A-34126	Sequence 34126, A
43	55	20.5	350	7	US-11-096-568A-34125	Sequence 34125, A
44	55	20.5	352	7	US-11-191-072-4	Sequence 4, Appli
45	55	20.5	523	7	US-11-087-099-7033	Sequence 7033, Ap

ALIGNMENTS

RESULT 1
US-11-169-041-167
; Sequence 167, Application US/11169041
; Publication No. US20060019284A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 10001 NP
; CURRENT APPLICATION NUMBER: US/11/169,041
; CURRENT FILING DATE: 2005-06-28
; PRIOR APPLICATION NUMBER: 60/584,405
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 527
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 167
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-169-041-167

Query Match 27.4%; Score 73.5; DB 7; Length 863;
Best Local Similarity 45.7%; Pred. No. 0.36;
Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

```
Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db        588 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 622
```

RESULT 2
US-11-096-568A-18567
; Sequence 18567, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 18567
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(219)
;   OTHER INFORMATION: Ceres Seq. ID no. 12365854
US-11-096-568A-18567

  Query Match             26.3%;  Score 70.5;  DB 7;  Length 219;
  Best Local Similarity   43.9%;  Pred. No. 0.18;
  Matches   18;  Conservative    2;  Mismatches   14;  Indels    7;  Gaps    1;

Qy          3 AGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKP 43
              |   | | |||| |       |  | : ||:||| ||   |
Db        185 ASGTPAAFCRAGTP-------RSSTPAPGSTSPATATTASP 218


RESULT 3
US-11-096-568A-3275
; Sequence 3275, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3275
;   LENGTH: 497
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(497)
;   OTHER INFORMATION: Ceres Seq. ID no. 15177393
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (312)..(312)
;   OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-3275

  Query Match             26.1%;  Score 70;  DB 7;  Length 497;
  Best Local Similarity   38.8%;  Pred. No. 0.52;
  Matches   19;  Conservative    9;  Mismatches    9;  Indels   12;  Gaps    3;

Qy          5 PRPGAPCRA--GAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSL 51
              | || | ::  | ||:        : :|| ||||:  |||  ::|| :|
Db        390 PIPGTPYKSAFGTPTL--------SPAHGISSPAK--LGKSPLINDSNL 428


RESULT 4
US-11-096-568A-3274
; Sequence 3274, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3274
;   LENGTH: 498
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(498)
;   OTHER INFORMATION: Ceres Seq. ID no. 15177392
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (313)..(313)
;   OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-3274

  Query Match             26.1%;  Score 70;  DB 7;  Length 498;
  Best Local Similarity   38.8%;  Pred. No. 0.52;
  Matches   19;  Conservative    9;  Mismatches    9;  Indels   12;  Gaps    3;

Qy          5 PRPGAPCRA--GAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSL 51
              | || | ::  | ||:        : :|| ||||:  |||  ::|| :|
Db        391 PIPGTPYKSAFGTPTL--------SPAHGISSPAK--LGKSPLINDSNL 429


RESULT 5
US-11-096-568A-3276
; Sequence 3276, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3276
;   LENGTH: 498
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(498)
;   OTHER INFORMATION: Ceres Seq. ID no. 16625597
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (313)..(313)
;   OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-3276

  Query Match             26.1%;  Score 70;  DB 7;  Length 498;
  Best Local Similarity   38.8%;  Pred. No. 0.52;
  Matches   19;  Conservative    9;  Mismatches    9;  Indels   12;  Gaps    3;

Qy          5 PRPGAPCRA--GAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSL 51
              | || | ::  | ||:        : :|| ||||:  |||  ::|| :|
Db        391 PIPGTPYKSAFGTPTL--------SPAHGISSPAK--LGKSPLINDSNL 429


RESULT 6
US-11-096-568A-3273
; Sequence 3273, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3273
;   LENGTH: 550
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(550)
;   OTHER INFORMATION: Ceres Seq. ID no. 15177391
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (365)..(365)
;   OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-3273
```

```
  Query Match             26.1%;  Score 70;  DB 7;  Length 550;
  Best Local Similarity   38.8%;  Pred. No. 0.58;
  Matches   19;  Conservative    9;  Mismatches    9;  Indels   12;  Gaps    3;

Qy          5 PRPGAPCRA--GAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLNDYSL 51
              | || | ::  | ||:       : :|| ||||:  |||  ::|| :|
Db        443 PIPGTPYKSAFGTPTL--------SPAHGISSPAK--LGKSPLINDSNL 481


RESULT 7
US-11-072-512-3112
; Sequence 3112, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3112
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-3112

  Query Match             23.5%;  Score 63;  DB 7;  Length 128;
  Best Local Similarity   35.6%;  Pred. No. 0.83;
  Matches   16;  Conservative    4;  Mismatches   23;  Indels    2;  Gaps    1;

Qy          4 GPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLND 48
              ||  |     | |    |:|    |  |:| | :  || || | :
Db          2 GPAAGHAAPRGPPVPATTAG--AALRAGASEPRQTQLGAPLALGE 44


RESULT 8
US-11-096-568A-12554
; Sequence 12554, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 12554
;   LENGTH: 251
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(251)
;   OTHER INFORMATION: Ceres Seq. ID no. 14302290
US-11-096-568A-12554

  Query Match             22.9%;  Score 61.5;  DB 7;  Length 251;
  Best Local Similarity   38.8%;  Pred. No. 2.8;
  Matches   19;  Conservative    2;  Mismatches   15;  Indels   13;  Gaps    2;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTLS----------HGSSSPARAT 39
              | || | || |   || : |:| |   |            ||  | |||
Db         28 ATGPSPSAPRR--TPTDLATAGSRSRRSPPTGPGTAPATGGSRRPRRAT 74


RESULT 9
US-11-096-568A-12552
; Sequence 12552, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 12552
;   LENGTH: 406
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(406)
;   OTHER INFORMATION: Ceres Seq. ID no. 14302288
US-11-096-568A-12552

  Query Match             22.9%;  Score 61.5;  DB 7;  Length 406;
  Best Local Similarity   38.8%;  Pred. No. 4.8;
  Matches   19;  Conservative    2;  Mismatches   15;  Indels   13;  Gaps    2;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTLS----------HGSSSPARAT 39
              | || | || |   || : |:| |   |            ||  | |||
Db        183 ATGPSPSAPRR--TPTDLATAGSRSRRSPPTGPGTAPATGGSRRPRRAT 229


RESULT 10
US-10-501-035-357
; Sequence 357, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
;  TITLE OF INVENTION:  ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE
;  TITLE OF INVENTION:  AND/OR PROTEIN TYROSINE KINASE PATHWAYS
;  FILE REFERENCE: D0185 PCT
;  CURRENT APPLICATION NUMBER: US/10/501,035
;  CURRENT FILING DATE:  2004-07-09
;  PRIOR APPLICATION NUMBER: US 60/350,061
;  PRIOR FILING DATE: 2002-01-18
;  NUMBER OF SEQ ID NOS: 795
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 357
;   LENGTH: 950
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-501-035-357

  Query Match             22.6%;  Score 60.5;  DB 6;  Length 950;
  Best Local Similarity   36.2%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches   16;  Indels   11;  Gaps    2;

Qy          3 AGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPAR-------ATLGK 42
```

```
              | ||   || |  |::|:  | |    |    | |        ||||
Db          10 ARPRSCPPCLAAGPSMVMADGPR----HLQRGPVRVGFYDIEGTLGK 52
```

RESULT 11
US-11-051-724-38
; Sequence 38, Application US/11051724
; Publication No. US20060063200A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjalt, Tord
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same,
; TITLE OF INVENTION: and Methods of Use
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/11/051,724
; CURRENT FILING DATE: 2005-02-02
; PRIOR APPLICATION NUMBER: US/10/162,335
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 38
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-051-724-38

Query Match 21.8%; Score 58.5; DB 7; Length 265;
Best Local Similarity 42.3%; Pred. No. 7;
Matches 22; Conservative 3; Mismatches 14; Indels 13; Gaps 4;

```
Qy          2 AAG------PRPGAPCRAGAPTIVLTSGRRQTLSHGSS--SPA-RATLGKPL 44
              |||      | || | |    ::   ||||     |:|  ||| | | | ||
```

```
Db        146 AAGGLVLPAPPPGRPVR----SVATQSGRRGGWGWGASAGSPASRGTGGNPL 193
```

RESULT 12
US-11-096-568A-23618
; Sequence 23618, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23618
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(375)
; OTHER INFORMATION: Ceres Seq. ID no. 12414168
US-11-096-568A-23618

Query Match 21.8%; Score 58.5; DB 7; Length 375;
Best Local Similarity 38.2%; Pred. No. 10;
Matches 13; Conservative 4; Mismatches 8; Indels 9; Gaps 1;

```
Qy          3 AGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPA 36
              ||||||  ||: : |         | :  :| ||
Db         80 AGPRPGGACRSSSTT---------TAAASASGPA 104
```

RESULT 13
US-11-229-769-199
; Sequence 199, Application US/11229769
; Publication No. US20060079670A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1D1C1
; CURRENT APPLICATION NUMBER: US/11/229,769
; CURRENT FILING DATE: 2005-09-20
; PRIOR APPLICATION NUMBER: 10/233,453
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 199
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-229-769-199

Query Match 21.6%; Score 58; DB 7; Length 117;
Best Local Similarity 30.4%; Pred. No. 3.2;
Matches 17; Conservative 4; Mismatches 19; Indels 16; Gaps 1;

```
Qy          1 MAAGPRPGA---------------PCRAGAPTIVLTSGRRQTLSHGSSSPARATL 40
```

```
              :|||| |                    |   ||  ||||  |: | | :  |  :|
Db         18 LAAGPTPSTGSSSPSWKQHIGTSLQKTRGSLPTTTLTSGAGQSTSTGKNPAAGRSL 73
```

RESULT 14
US-11-229-769-351
; Sequence 351, Application US/11229769
; Publication No. US20060079670A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1D1C1
; CURRENT APPLICATION NUMBER: US/11/229,769
; CURRENT FILING DATE: 2005-09-20
; PRIOR APPLICATION NUMBER: 10/233,453
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 351
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-229-769-351

Query Match 21.6%; Score 58; DB 7; Length 127;
Best Local Similarity 30.4%; Pred. No. 3.5;
Matches 17; Conservative 4; Mismatches 19; Indels 16; Gaps 1;

```
Qy          1 MAAGPRPGA----------------PCRAGAPTIVLTSGRRQTLSHGSSSPARATL 40
              :|||| |                  |   ||  ||||  |: | | :  |  :|
Db         28 LAAGPTPSTGSSSPSWKQHIGTSLQKTRGSLPTTTLTSGAGQSTSTGKNPAAGRSL 83
```

RESULT 15
US-11-229-769-238
; Sequence 238, Application US/11229769
; Publication No. US20060079670A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1D1C1
; CURRENT APPLICATION NUMBER: US/11/229,769
; CURRENT FILING DATE: 2005-09-20
; PRIOR APPLICATION NUMBER: 10/233,453
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 238
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-229-769-238

Query Match 21.6%; Score 58; DB 7; Length 225;
Best Local Similarity 30.4%; Pred. No. 6.7;
Matches 17; Conservative 4; Mismatches 19; Indels 16; Gaps 1;

```
Qy          1 MAAGPRPGA----------------PCRAGAPTIVLTSGRRQTLSHGSSSPARATL 40
              :|||| |                  |   ||  ||||  |: | | :  |  :|
Db        126 LAAGPTPSTGSSSPSWKQHIGTSLQKTRGSLPTTTLTSGAGQSTSTGKNPAAGRSL 181
```

Search completed: April 27, 2006, 15:41:21
Job time : 27 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 27, 2006, 15:15:05 ; Search time 36.7059 Seconds
(without alignments)
136.307 Million cell updates/sec

Title: US-09-819-144A-2
Perfect score: 268
Sequence: 1 MAAGPRPGAPCRAGAPTIVL.....SSPARATLGKPLVLNDYSLN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	68	25.4	315	2	C70810	hypothetical prote
2	60.5	22.6	504	2	T33404	hypothetical prote
3	60	22.4	650	2	T48060	beta-D-glucan exoh
4	59	22.0	1687	2	T30176	EGF repeat transme
5	59	22.0	4180	2	G83559	hypothetical prote
6	58.5	21.8	323	2	T19142	hypothetical prote
7	58	21.6	419	2	T36272	hypothetical prote
8	57.5	21.5	886	2	S07132	hypothetical prote
9	57.5	21.5	1611	2	T38236	hypothetical prote
10	57	21.3	308	2	D70875	probable PE protei
11	57	21.3	313	2	A46233	FLAT element-bindi
12	57	21.3	457	2	C70589	probable cobB prot
13	57	21.3	474	2	H86303	hypothetical prote
14	57	21.3	775	2	D86261	hypothetical prote
15	57	21.3	1487	1	EDBEE1	immediate-early pr
16	57	21.3	1487	1	EDBEF6	155K transcription
17	56.5	21.1	136	2	T36547	hypothetical prote
18	56.5	21.1	205	2	C89957	hypothetical prote
19	56.5	21.1	408	2	B87436	conserved hypothet
20	56.5	21.1	640	2	A87715	glucose inhibited
21	56.5	21.1	757	2	T09081	telomere-associate
22	56	20.9	232	2	S24390	transforming prote
23	56	20.9	430	2	T46420	hypothetical prote
24	56	20.9	540	2	S21825	vicilin-like stora
25	56	20.9	616	2	A72627	probable 2-oxoacid
26	56	20.9	780	2	F84470	probable retroelem
27	55.5	20.7	701	2	S61239	hypothetical prote
28	55.5	20.7	1067	2	D75625	probable extracell
29	55.5	20.7	1408	2	S16148	gene serrate prote
30	55.5	20.7	1724	2	T18343	P-glycoprotein - S
31	55	20.5	296	2	A40996	phenylalanine 4-mo
32	55	20.5	323	2	C86384	unknown protein [i
33	55	20.5	352	2	JC2466	inhibin beta-C cha
34	55	20.5	364	1	TVHUML	transforming prote
35	55	20.5	445	2	E87561	hypothetical prote
36	55	20.5	493	2	G84263	long-chain fatty-a
37	55	20.5	513	2	AD1959	ATP-dependent RNA
38	55	20.5	677	2	E70722	hypothetical prote
39	54.5	20.3	251	2	A55523	hypothetical prote
40	54.5	20.3	420	2	AG1859	hypothetical prote
41	54.5	20.3	552	2	E87226	conserved membrane
42	54.5	20.3	848	2	D85635	trimethylamine N-o
43	54.5	20.3	848	2	H90772	trimethylamine N-o
44	54.5	20.3	848	2	C64841	trimethylamine-N-o
45	54	20.1	178	2	T30717	probable virion pr

ALIGNMENTS

RESULT 1
C70810
hypothetical protein Rv0819 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: C70810
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70810
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-315 <COL>
A;Cross-references: UNIPROT:O53831; UNIPARC:UPI00000D119B; GB:AL022004; GB:AL123456; NI
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv0819
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0819

Query Match 25.4%; Score 68; DB 2; Length 315;
Best Local Similarity 28.2%; Pred. No. 1.9;
Matches 20; Conservative 5; Mismatches 14; Indels 32; Gaps 2;

```
Qy          1 MAAGPRPGAPC-------------------------RAGAPT------IVLTSGRRQTL 28
              : || ||| |                           | |  |      :  |:|| |
Db         51 LVAGSRPGGPIIGYLNLSPPRGAGGAMAELVVHPQSRRRGIGTAMARAALAKTAGRNQFW 110

Qy         29 SHGSSSPARAT 39
                :||:  |||||
Db        111 AHGTLDPARAT 121
```

RESULT 2
T33404
hypothetical protein H10E21.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33404
R;Davidson, S.; Wohldmann, P.; Courtney, L.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of C. elegans cosmid H10E21.
A;Reference number: Z21339
A;Accession: T33404
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-504 <DAV>
A;Cross-references: UNIPROT:Q8ITW8; UNIPARC:UPI0000084101; EMBL:AF078783; PIDN:AAC26921
A;Experimental source: strain Bristol N2; clone H10E21

```
C;Genetics:
A;Gene: CESP:H10E21.3
A;Map position: 3
A;Introns: 78/1; 168/3; 241/3; 384/1; 434/3

  Query Match             22.6%;  Score 60.5;  DB 2;  Length 504;
  Best Local Similarity   29.2%;  Pred. No. 23;
  Matches   19;  Conservative    5;  Mismatches   24;  Indels   17;  Gaps    2;

Qy          4 GPRPGAPCRAGAPTIVL------TSGRRQ-----------TLSHGSSSPARATLGKPLVL 46
              || | ||| ||   : |         | |:           |      | :    ||||
Db        194 GPSPMAPCSAGPDVLTLFIREEMKLGERRRLLFSERAVGTLLGQNKHCPYKKEDIKPLVF 253

Qy         47 NDYSL 51
              :|: |
Db        254 HDFRL 258


RESULT 3
T48060
beta-D-glucan exohydrolase-like protein - Arabidopsis thaliana
N;Alternate names: protein F26K9.140
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 31-Dec-2004
C;Accession: T48060
R;Bloecker, H.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanouba
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24465
A;Accession: T48060
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-650 <BLO>
A;Cross-references: UNIPROT:Q9LZJ4; UNIPARC:UPI00000488BE; EMBL:AL162651
A;Experimental source: cultivar Columbia; BAC clone F26K9
C;Genetics:
A;Map position: 3
A;Introns: 94/1; 159/2; 189/3; 232/2; 348/3; 410/3; 493/1
A;Note: F26K9.140

  Query Match             22.4%;  Score 60;  DB 2;  Length 650;
  Best Local Similarity   30.2%;  Pred. No. 34;
  Matches   16;  Conservative    8;  Mismatches   25;  Indels    4;  Gaps    1;

Qy          1 MAAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATL----GKPLVLNDY 49
              :  |  | |     :||: :|     ||||   |  :  :    |:|||:  |
Db        532 VVVGETPYAETFGDSPTLGITKPGPDTLSHTCGSGMKCLVILVTGRPLVIEPY 584


RESULT 4
T30176
EGF repeat transmembrane protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30176
R;Sell, C.; Hoff III, H.B.
submitted to the EMBL Data Library, May 1996
A;Description: Cloning of a novel mRNA regulated by the insulin like growth factor type
A;Reference number: Z20762
A;Accession: T30176
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1687 <SEL>
A;Cross-references: UNIPROT:Q61204; UNIPARC:UPI00000280BB; EMBL:U57368; NID:g1336627; PI
A;Experimental source: strain C57BL/6J; clone DBI-1; whole embryo

  Query Match             22.0%;  Score 59;  DB 2;  Length 1687;
  Best Local Similarity   37.1%;  Pred. No. 1.2e+02;
  Matches   26;  Conservative    3;  Mismatches   21;  Indels   20;  Gaps    5;

Qy          3 AGP-----RPGAPCRAGAP-----TIVLTSGRRQT---LSH--GSSSPARAT-----LGK 42
              |||     ||| |   | |         |  ||||:    ||  |   ||  |     | |
Db       1439 AGPQNKHKRPGEPSMQGIPKRRRCASPLLRGRRQSPAVNSHIGGKGPPAPMTQAQPGLIK 1498

Qy         43 PLVLNDYSLN 52
              || |:  : |
Db       1499 PLPLHKEATN 1508


RESULT 5
G83559
hypothetical protein PA0690 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: G83559
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: G83559
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-4180 <STO>
A;Cross-references: UNIPROT:Q9I5N6; UNIPARC:UPI0000110216; GB:AE004504; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA0690

  Query Match             22.0%;  Score 59;  DB 2;  Length 4180;
  Best Local Similarity   41.4%;  Pred. No. 2.9e+02;
  Matches   12;  Conservative    6;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         13 AGAPTIVLTSGRRQTLSHGSSSPARATLG 41
              :|: |:||  | |  | ||:::||    |
Db       2565 SGSGTLVLAGGARIDLRHGTAAPAEQVDG 2593


RESULT 6
T19142
hypothetical protein C09G5.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19142
R;Palmer, S.
submitted to the EMBL Data Library, November 1994
A;Reference number: Z19080
A;Accession: T19142
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-323 <WIL>
A;Cross-references: UNIPROT:Q09455; UNIPARC:UPI0000127242; EMBL:Z46791; PIDN:CAA86757.1;
A;Experimental source: clone C09G5
C;Genetics:
A;Gene: CESP:C09G5.4
A;Map position: 2
A;Introns: 47/3

  Query Match             21.8%;  Score 58.5;  DB 2;  Length 323;
  Best Local Similarity   41.5%;  Pred. No. 25;
  Matches   17;  Conservative    3;  Mismatches   16;  Indels    5;  Gaps    1;

Qy          3 AGPR-----PGAPCRAGAPTIVLTSGRRQTLSHGSSSPARA 38
              ||||     ||| : |||     :||    |||   |: |
Db        170 AGPRGPAGDAGAPGQVGAPGNPGQAGRGGQRSHGLPGPSGA 210


RESULT 7
T36272
hypothetical protein SCE68.22 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36272
R;Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A;Reference number: Z21576
```

```
A;Accession: T36272
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-419 <MUR>
A;Cross-references: UNIPROT:Q9WX10; UNIPARC:UPI00000DB13A; EMBL:AL079345; PIDN:CAB45358.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCE68.22

  Query Match             21.6%;  Score 58;  DB 2;  Length 419;
  Best Local Similarity   38.5%;  Pred. No. 37;
  Matches   15;  Conservative    3;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATL 40
              ||  : | |  |||       ||  ||:||  :  |  |
Db         29 AAAEQSGGPATAGATAEPTEQGRLVTLAHGLETLPRPEL 67


RESULT 8
S07132
hypothetical protein 1 - Rhodospirillum rubrum (fragment)
C;Species: Rhodospirillum rubrum
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: S07132
R;Falk, G.; Hampe, A.; Walker, J.E.
Biochem. J. 228, 391-407, 1985
A;Title: Nucleotide sequence of the Rhodospirillum rubrum atp operon.
A;Reference number: S07132; MUID:85251588; PMID:2861810
A;Accession: S07132
A;Molecule type: DNA
A;Residues: 1-886 <FAL>
A;Cross-references: UNIPROT:Q9LCV5; UNIPARC:UPI00000AFF84; EMBL:X02499; NID:g46360; PIDN

  Query Match             21.5%;  Score 57.5;  DB 2;  Length 886;
  Best Local Similarity   27.9%;  Pred. No. 91;
  Matches   17;  Conservative    3;  Mismatches   22;  Indels   19;  Gaps    1;

Qy          3 AGPRP------------------GAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKP 43
              | |||                  | |  ||   ||:  |:||   ||          |:
Db        513 ARPRPPNHADFLPSPQRHPHQGTDGGPGHAGGHLIVIGPGQRQRQQHGDDGLGVGVFGEK 572

Qy         44 L 44
              |
Db        573 L 573


RESULT 9
T38236
hypothetical protein SPAC23A1.17 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T38236
R;Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21780
A;Accession: T38236
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1611 <MUR>
A;Cross-references: UNIPROT:O42854; UNIPARC:UPI0000069573; EMBL:AL021813; PIDN:CAA16991.
A;Experimental source: strain 972h-; cosmid c23A1
C;Genetics:
A;Gene: SPDB:SPAC23A1.17
A;Map position: 1

  Query Match             21.5%;  Score 57.5;  DB 2;  Length 1611;
  Best Local Similarity   35.0%;  Pred. No. 1.7e+02;
  Matches   14;  Conservative    8;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          5 PRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPL 44
              |  |:  |   |  ::||| |  |    :||: :::| ||
Db        801 PSTGSQLRNVVPVSIVTSGGRPALPDEMASPS-SSIGHPL 839


RESULT 10
D70875
probable PE protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: D70875
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70875
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-308 <COL>
A;Cross-references: UNIPROT:O50428; UNIPARC:UPI00000D5E5F; GB:AL010186; GB:AL123456; NID
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: PE

  Query Match             21.3%;  Score 57;  DB 2;  Length 308;
  Best Local Similarity   48.0%;  Pred. No. 36;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         26 QTLSHGSSSPARATLGKPLVLNDYS 50
              ||: :  ::|||| || ||:  | |
Db         97 QTMVNAVNAPARALLGHPLISADAS 121


RESULT 11
A46233
FLAT element-binding basic helix-loop-helix protein cdx-3 - golden hamster
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 31-Dec-2004
C;Accession: A46233
R;German, M.S.; Wang, J.; Chadwick, R.B.; Rutter, W.J.
Genes Dev. 6, 2165-2176, 1992
A;Title: Synergistic activation of the insulin gene by a LIM-homeo domain protein and a
A;Reference number: A46233; MUID:93051335; PMID:1358758
A;Accession: A46233
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-313 <GER>
A;Cross-references: UNIPROT:Q04649; UNIPARC:UPI0000127472; GB:X81404; GB:X69557; NID:g58
A;Experimental source: insulinoma cell line HIT T-15 M2.2.2
A;Note: sequence extracted from NCBI backbone (NCBIN:117925, NCBIP:117926)
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;186-242/Domain: homeobox homology <HOX>

  Query Match             21.3%;  Score 57;  DB 2;  Length 313;
  Best Local Similarity   35.1%;  Pred. No. 36;
  Matches   13;  Conservative    5;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          5 PRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLG 41
              | |  |   ||   :  ||:|: |       ||: :||
Db        141 PGPPGPAATGAAEQLSPSGQRRNLCEWMRKPAQPSLG 177


RESULT 12
C70589
probable cobB protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: C70589
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
```

```
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70589
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-457 <COL>
A;Cross-references: UNIPROT:O05811; UNIPARC:UPI0000127DCA; GB:Z95207; GB:AL123456; NID:g
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: cobB
C;Superfamily: cobyrinic acid a,c-diamide synthase

  Query Match             21.3%;  Score 57;  DB 2;  Length 457;
  Best Local Similarity   32.8%;  Pred. No. 53;
  Matches   19;  Conservative    6;  Mismatches   19;  Indels   14;  Gaps    2;

Qy          4 GPRPGAPCRA---------GAPTIVLTSGRRQT-----LSHGSSSPARATLGKPLVLN 47
              ||  |||             ||| |::   | |:     | || |:   ||    ::||
Db         99 GPAGGAPAAGSTAHVAALLGAPVILVVDARGQSHSVAALLHGFSTFDTATRIAGVILN 156


RESULT 13
H86303
hypothetical protein F6I1.14 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: H86303
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H86303
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-474 <STO>
A;Cross-references: UNIPROT:Q9FZ45; UNIPARC:UPI00000AC1A8; GB:AE005172; NID:g9802778; PI
C;Genetics:
A;Map position: 1

  Query Match             21.3%;  Score 57;  DB 2;  Length 474;
  Best Local Similarity   35.9%;  Pred. No. 55;
  Matches   14;  Conservative    7;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          3 AGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLG 41
              :|| |||| |:|:      ||    :: |  | : |: |
Db         64 SGPIPGAPSRSGSFAGTAQSGPGAPMATGRMSGSLASAG 102


RESULT 14
D86261
hypothetical protein F13K23.5 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D86261
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86261
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-775 <STO>
A;Cross-references: UNIPROT:Q9LPW9; UNIPARC:UPI00000A425F; GB:AE005172; NID:g8698727; PII
C;Genetics:
A;Map position: 1

  Query Match             21.3%;  Score 57;  DB 2;  Length 775;
  Best Local Similarity   35.6%;  Pred. No. 91;
  Matches   16;  Conservative    7;  Mismatches   20;  Indels    2;  Gaps    1;

Qy          5 PRPGAPCRAGAPTIVLTS--GRRQTLSHGSSSPARATLGKPLVLN 47
              |    | :  : |||  |  |: | |    |: |:   :||| |:|
Db        346 PLERIPSKPISQTIVEASLQGKPQRLDPSSAEPSVPNIGKPSVVN 390


RESULT 15
EDBEE1
immediate-early protein - equine herpesvirus 1
C;Species: equine herpesvirus 1
C;Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 09-Jul-2004
C;Accession: A33764
R;Grundy, F.J.; Baumann, R.P.; O'Callaghan, D.J.
Virology 172, 223-236, 1989
A;Title: DNA sequence and comparative analyses of the equine herpesvirus type 1 immediate
A;Reference number: A33764; MUID:89370304; PMID:2549711
A;Accession: A33764
A;Molecule type: DNA
A;Residues: 1-1487 <GRU>
A;Cross-references: UNIPROT:P17473; UNIPARC:UPI000004C6C8; GB:J04366; NID:g330910; PIDN:A
C;Superfamily: herpesvirus immediate-early protein IE175
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             21.3%;  Score 57;  DB 1;  Length 1487;
  Best Local Similarity   29.8%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    8;  Mismatches   18;  Indels   14;  Gaps    2;

Qy          2 AAGPRPGAPCRAGAP----------TIVLTSGRRQTLSHGSSSPARATLGKPLVLN 47
              ||  :| :|  : ||           ::   :||| |   |||||       |: |:
Db         97 AAAAQPASPAPSPAPGLAAMLKMVHSSVAPGNGRRAT---GSSSPGGGDAADPVALD 150


Search completed: April 27, 2006, 15:20:32
Job time : 38.7059 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 27, 2006, 15:09:59 ; Search time 224.824 Seconds
(without alignments)
163.183 Million cell updates/sec

Title: US-09-819-144A-2
Perfect score: 268
Sequence: 1 MAAGPRPGAPCRAGAPTIVL.....SSPARATLGKPLVLNDYSLN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	73.5	27.4	290	2	Q68D58_HUMAN	Q68d58 homo sapien
2	73.5	27.4	350	2	Q641Q8_HUMAN	Q641q8 homo sapien
3	73.5	27.4	863	1	MILK1_HUMAN	Q8n3f8 homo sapien
4	72	26.9	636	2	Q6AYD5_RAT	Q6ayd5 rattus norv
5	71.5	26.7	1581	2	Q4SA50_TETNG	Q4sa50 tetraodon n
6	70	26.1	1004	2	Q4P5V3_USTMA	Q4p5v3 ustilago ma
7	68	25.4	315	2	Q7U173_MYCBO	Q7u173 mycobacteri
8	68	25.4	315	2	O53831_MYCTU	O53831 mycobacteri
9	67.5	25.2	391	2	Q98A00_RHILO	Q98a00 rhizobium l
10	67	25.0	563	2	Q4UVP4_XANCP	Q4uvp4 xanthomonas
11	67	25.0	563	2	Q8P8E3_XANCP	Q8p8e3 xanthomonas
12	66.5	24.8	1787	2	Q9M4X9_CHLRE	Q9m4x9 chlamydomon
13	66	24.6	272	2	Q743X3_MYCPA	Q743x3 mycobacteri
14	64.5	24.1	215	2	Q8IV50_HUMAN	Q8iv50 homo sapien
15	64.5	24.1	596	2	Q82NB1_STRAW	Q82nb1 streptomyce
16	64	23.9	613	2	Q4WQ64_ASPFU	Q4wq64 aspergillus
17	64	23.9	731	2	Q4Q5C4_LEIMA	Q4q5c4 leishmania
18	63.5	23.7	355	2	Q5R9T2_PONPY	Q5r9t2 pongo pygma
19	63.5	23.7	377	2	Q8N3K4_HUMAN	Q8n3k4 homo sapien
20	63.5	23.7	562	2	Q8BQZ7_MOUSE	Q8bqz7 mus musculu
21	63.5	23.7	874	2	Q6PDL7_MOUSE	Q6pdl7 mus musculu
22	63	23.5	128	2	Q8N0Y6_HUMAN	Q8n0y6 homo sapien
23	63	23.5	171	2	Q6K3G1_ORYSA	Q6k3g1 oryza sativ
24	63	23.5	206	2	Q14897_HUMAN	Q14897 homo sapien
25	63	23.5	236	2	Q5QPL0_HUMAN	Q5qpl0 homo sapien
26	63	23.5	360	1	YBOX2_MOUSE	Q9z2c8 mus musculu
27	63	23.5	373	2	Q7WNE3_BORBR	Q7wne3 bordetella
28	63	23.5	2063	2	Q4QG78_LEIMA	Q4qg78 leishmania
29	62.5	23.3	150	2	Q84ZQ6_ORYSA	Q84zq6 oryza sativ
30	62.5	23.3	208	2	Q5Z9G0_ORYSA	Q5z9g0 oryza sativ
31	62.5	23.3	556	2	Q8XQ41_RALSO	Q8xq41 ralstonia s
32	62.5	23.3	671	2	Q9Z5X4_FRASE	Q9z5x4 frankia sp.
33	62.5	23.3	771	2	Q5ZKM6_CHICK	Q5zkm6 gallus gall
34	62	23.1	215	2	O93392_COTCO	O93392 coturnix co
35	62	23.1	259	2	Q6ZMZ8_HUMAN	Q6zmz8 homo sapien
36	62	23.1	415	2	Q6A047_MOUSE	Q6a047 mus musculu
37	62	23.1	449	2	Q7SG58_NEUCR	Q7sg58 neurospora
38	62	23.1	496	2	Q7U8L8_SYNPX	Q7u8l8 synechococc
39	62	23.1	1050	2	Q4T9N9_TETNG	Q4t9n9 tetraodon n
40	62	23.1	1059	2	Q5KG17_CRYNE	Q5kg17 cryptococcu
41	62	23.1	1059	2	Q55S85_CRYNE	Q55s85 cryptococcu
42	62	23.1	1325	1	RIMB2_CHICK	Q8qfx1 gallus gall
43	62	23.1	1553	2	Q4SE53_TETNG	Q4se53 tetraodon n
44	61.5	22.9	588	2	Q9N2G7_RABIT	Q9n2g7 oryctolagus
45	61.5	22.9	1178	2	Q69ZH8_MOUSE	Q69zh8 mus musculu

ALIGNMENTS

RESULT 1
Q68D58_HUMAN
ID Q68D58_HUMAN PRELIMINARY; PRT; 290 AA.
AC Q68D58;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp686M2226.
GN Name=DKFZp686M2226;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RG The German cDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; CR749566; CAH18362.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 290 AA; 32971 MW; DF4D273FB8C84BCB CRC64;

Query Match 27.4%; Score 73.5; DB 2; Length 290;
Best Local Similarity 45.7%; Pred. No. 2.5;
Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

```
Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:| ||    || :| | | : | |||||
Db         15 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 49
```

RESULT 2
Q641Q8_HUMAN
ID Q641Q8_HUMAN PRELIMINARY; PRT; 350 AA.
AC Q641Q8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MICAL-L1 protein (Fragment).
GN Name=MICAL-L1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

```
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Eye;
RG   NIH MGC Project;
RL   Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC082243; AAH82243.1; -; mRNA.
FT   NON_TER       1       1
SQ   SEQUENCE   350 AA;  38893 MW;  8F1600A6CFC6345A CRC64;

  Query Match             27.4%;  Score 73.5;  DB 2;  Length 350;
  Best Local Similarity   45.7%;  Pred. No. 3.1;
  Matches   16;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db         75 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 109


RESULT 3
MILK1_HUMAN
ID   MILK1_HUMAN     STANDARD;      PRT;   863 AA.
AC   Q8N3F8; Q5TI16; Q7RTP5; Q8N3N8; Q9BVL9; Q9BY92; Q9UH43; Q9UH44;
AC   Q9UH45;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Molecule interacting with Rab13 (MIRab13) (MICAL-like protein 1).
GN   Name=MIRAB13; Synonyms=KIAA1668;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Uterine adenocarcinoma;
RA   Pandjaitan R., Zahraoui A.;
RT   "MIRab13, a novel effector of Rab13.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX   PubMed=15461802; DOI=10.1186/gb-2004-5-10-r84;
RA   Collins J.E., Wright C.L., Edwards C.A., Davis M.P., Grinham J.A.,
RA   Cole C.G., Goward M.E., Aguado B., Mallya M., Mokrab Y., Huckle E.J.,
RA   Beare D.M., Dunham I.;
RT   "A genome annotation-driven approach to cloning the human ORFeome.";
RL   Genome Biol. 5:RESEARCH84.1-RESEARCH84.11(2004).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;
RA   Dunham I., Hunt A.R., Collins J.E., Bruskiewich R., Beare D.M.,
RA   Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA   Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA   Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA   Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA   Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA   Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA   Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA   Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA   Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA   Hall R.E., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,
RA   Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA   Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA   Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA   Mcclay J., Mclaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA   Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA   Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA   Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA   Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA   Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA   Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA   Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA   Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA   Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA   Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA   Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA   Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA   Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA   Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA   Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
RA   Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA   Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA   Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA   Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,
RA   Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
RA   Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA   Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA   Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
RA   Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
RA   O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA   Khan A.S., Lane L., Tilahun Y., Wright H.;
RT   "The DNA sequence of human chromosome 22.";
RL   Nature 402:489-495(1999).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 25-863, AND VARIANT SER-519.
RC   TISSUE=Melanoma;
RG   The German cDNA consortium;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 73-863.
RC   TISSUE=Brain;
RX   MEDLINE=21156230; PubMed=11258795;
RA   Hirosawa M., Nagase T., Murahashi Y., Kikuno R., Ohara O.;
RT   "Identification of novel transcribed sequences on human chromosome 22
RT   by expressed sequence tag mapping.";
RL   DNA Res. 8:1-9(2001).
RN   [6]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 514-863.
RC   TISSUE=Choriocarcinoma;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
```

```
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   GENE STRUCTURE.
RX   MEDLINE=22105479; PubMed=12110185; DOI=10.1016/S0092-8674(02)00794-8;
RA   Terman J.R., Mao T., Pasterkamp R.J., Yu H.-H., Kolodkin A.L.;
RT   "MICALs, a family of conserved flavoprotein oxidoreductases, function
RT   in plexin-mediated axonal repulsion.";
RL   Cell 109:887-900(2002).
RN   [8]
RP   PHOSPHORYLATION SITE SER-578.
RX   PubMed=15302935; DOI=10.1073/pnas.0404720101;
RA   Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J.,
RA   Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
RT   "Large-scale characterization of HeLa cell nuclear phosphoproteins.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135(2004).
CC   -!- FUNCTION: May be a cytoskeletal regulator.
CC   -!- SUBUNIT: Binds to Rab13.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Contains 1 CH (calponin-homology) domain.
CC   -!- SIMILARITY: Contains 1 LIM zinc-binding domain.
CC   -!- CAUTION: Ref.4 (CAD39036) sequence differs from that shown due to
CC       frameshifts in positions 486 and 507.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AJ496196; CAD42713.1; -; mRNA.
DR   EMBL; CR456437; CAG30323.1; -; mRNA.
DR   EMBL; AL022311; CAI18864.1; -; Genomic_DNA.
DR   EMBL; AL834373; CAD39036.1; ALT_FRAME; mRNA.
DR   EMBL; AL833860; CAD38718.1; -; mRNA.
DR   EMBL; AB051455; BAB33338.1; -; mRNA.
DR   EMBL; BC001090; AAH01090.2; -; mRNA.
DR   EMBL; BK000466; DAA01345.1; -; mRNA.
DR   HSSP; Q01082; 1BKR.
DR   Ensembl; ENSG00000100139; Homo sapiens.
DR   GO; GO:0005622; C:intracellular; IC.
DR   GO; GO:0016192; P:vesicle-mediated transport; IC.
DR   InterPro; IPR001715; Calponin_act_bd.
DR   InterPro; IPR001781; LIM_Zn_bd.
DR   Pfam; PF00307; CH; 1.
DR   Pfam; PF00412; LIM; 1.
DR   ProDom; PD000094; LIM; 1.
DR   SMART; SM00033; CH; 1.
DR   SMART; SM00132; LIM; 1.
DR   PROSITE; PS50021; CH; 1.
DR   PROSITE; PS00478; LIM_DOMAIN_1; 1.
DR   PROSITE; PS50023; LIM_DOMAIN_2; 1.
KW   Coiled coil; Cytoskeleton; LIM domain; Metal-binding; Phosphorylation;
KW   Polymorphism; Zinc.
FT   DOMAIN        2    105       CH.
FT   DOMAIN      162    225       LIM zinc-binding.
FT   COILED      682    711       Potential.
FT   COILED      785    830       Potential.
FT   COMPBIAS    253    667       Pro-rich.
FT   COMPBIAS    428    437       Poly-Glu.
FT   MOD_RES     578    578       Phosphoserine.
FT   VARIANT     519    519       A -> S (in dbSNP:9610875).
FT                                /FTId=VAR_018262.
FT   VARIANT     583    583       P -> L (in dbSNP:2272829).
FT                                /FTId=VAR_020258.
FT   CONFLICT    137    137       P -> S (in Ref. 5).
FT   CONFLICT    212    212       F -> Y (in Ref. 4; CAD39036).
SQ   SEQUENCE   863 AA;  93441 MW;  755E3B57C6037292 CRC64;

  Query Match            27.4%;  Score 73.5;  DB 1;  Length 863;
  Best Local Similarity  45.7%;  Pred. No. 8.7;
  Matches   16;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
              ::||:|  ||    || :|  | |  : | |||||
Db        588 SSGPQPAKPCSGATPTPLLLVGDRSPVPSPGSSSP 622


RESULT 4
Q6AYD5_RAT
ID   Q6AYD5_RAT  PRELIMINARY;      PRT;   636 AA.
AC   Q6AYD5;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   G1 to S phase transition 1.
GN   Name=Gspt1;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Lung;
RA   Director MGC Project;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC079092; AAH79092.1; -; mRNA.
DR   GO; GO:0005622; C:intracellular; ISS.
DR   GO; GO:0003924; F:GTPase activity; ISS.
DR   GO; GO:0005515; F:protein binding; ISS.
DR   GO; GO:0003747; F:translation release factor activity; ISS.
DR   GO; GO:0000082; P:G1/S transition of mitotic cell cycle; ISS.
DR   GO; GO:0000184; P:mRNA catabolism, nonsense-mediated decay; ISS.
DR   InterPro; IPR009818; Ataxin-2_C.
DR   InterPro; IPR004160; EFTU_Cterm.
DR   InterPro; IPR004161; EFTU_D2.
DR   InterPro; IPR000795; ProtSyn_GTPbind.
DR   Pfam; PF00009; GTP_EFTU; 1.
DR   Pfam; PF03144; GTP_EFTU_D2; 1.
DR   Pfam; PF03143; GTP_EFTU_D3; 1.
DR   Pfam; PF07145; PAM2; 1.
DR   PRINTS; PR00315; ELONGATNFCT.
DR   PROSITE; PS00301; EFACTOR_GTP; UNKNOWN_1.
SQ   SEQUENCE   636 AA;  68751 MW;  6FB5F124F5132D2C CRC64;

  Query Match            26.9%;  Score 72;  DB 2;  Length 636;
  Best Local Similarity  39.1%;  Pred. No. 9.3;
  Matches   18;  Conservative    7;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          2 AAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLN 47
              |||   ||   || |:  : | : |:|  ||:|     | :|:| |
```

```
Db        101 AAGGDHGAGSSAGGPSEPVESSQEQSLCEGSNSTVSMELSEPVVEN 146


RESULT 5
Q4SA50_TETNG
ID   Q4SA50_TETNG PRELIMINARY;      PRT;  1581 AA.
AC   Q4SA50;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 12 SCAF14692, whole genome shotgun sequence.
GN   ORFNames=GSTENG00021607001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR   EMBL; CAAE01014692; CAG02482.1; -; Genomic_DNA.
DR   InterPro; IPR007108; Cut_homeo.
DR   InterPro; IPR003350; Hmoeo_CUT.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF02376; CUT; 3.
DR   Pfam; PF00046; Homeobox; 1.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS51042; CUT; 3.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
SQ   SEQUENCE   1581 AA;  170831 MW;  710958B08CC7597A CRC64;

  Query Match             26.7%;  Score 71.5;  DB 2;  Length 1581;
  Best Local Similarity   44.2%;  Pred. No. 30;
  Matches   19;  Conservative    9;  Mismatches   12;  Indels    3;  Gaps    2;

Qy          1 MAAGPRPGA--PCRAGAPTIVLTSGRRQTLSHGS-SSPARATL 40
              | : |:|||  | :||:|: | : |   : :  | |:||||||
Db       1065 MLSRPKPGASSPRKAGSPSSVCSCGYWTSSARASASTPARATL 1107


RESULT 6
Q4P5V3_USTMA
ID   Q4P5V3_USTMA PRELIMINARY;      PRT;  1004 AA.
AC   Q4P5V3;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=UM04510.1;
OS   Ustilago maydis 521.
OC   Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC   Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX   NCBI_TaxID=237631;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=521;
RA   Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA   Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA   Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA   Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA   Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA   Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA   Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA   David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA   Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA   Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA   Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA   Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA   Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA   Jaffe D., Jones C., Kamal M., Kamat A., Kamvysselis M., Karlsson E.,
RA   Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA   Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA   Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA   Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA   Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA   Mccarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA   Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA   Mozes J., Mulrain L., Munson G., Naylor J., Newes C., Nguyen C.,
RA   Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA   Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA   O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piqani B.,
RA   Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA   Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA   Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA   Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA   Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA   Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA   Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA   Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA   Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA   Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA   Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA   Zimmer A., Zody M., Lander E.;
RT   "The genome sequence of Ustilago maydis.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACP01000160; EAK85392.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   1004 AA;  108726 MW;  3948B22248CF953A CRC64;

  Query Match             26.1%;  Score 70;  DB 2;  Length 1004;
  Best Local Similarity   42.6%;  Pred. No. 27;
  Matches   20;  Conservative    2;  Mismatches   21;  Indels    4;  Gaps    1;

Qy          3 AGPRPGAPCRAGAPTIVLT----SGRRQTLSHGSSSPARATLGKPLV 45
              | |||   | |   |  ||    : || :||   ||| | |   | |
Db        539 AAPRPNVDCEALPQTAPLTLTKQATRRSSLSKDDSSPPRPTRPAPPV 585


RESULT 7
Q7U173_MYCBO
ID   Q7U173_MYCBO PRELIMINARY;      PRT;   315 AA.
AC   Q7U173;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein Mb0842.
GN   OrderedLocusNames=Mb0842;
```

```
OS   Mycobacterium bovis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium tuberculosis complex.
OX   NCBI_TaxID=1765;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AF2122/97;
RX   MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA   Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA   Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA   Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA   Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT   "The complete genome sequence of Mycobacterium bovis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR   EMBL; BX248336; CAD93704.1; -; Genomic_DNA.
DR   HSSP; O53831; 1P0H.
DR   SMR; Q7U173; 3-310.
DR   GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR   InterPro; IPR000182; GCN5acetyl_trans.
DR   Pfam; PF00583; Acetyltransf_1; 2.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   315 AA;  33599 MW;  E1728C0652CF2716 CRC64;

  Query Match             25.4%;  Score 68;  DB 2;  Length 315;
  Best Local Similarity   28.2%;  Pred. No. 12;
  Matches   20;  Conservative    5;  Mismatches   14;  Indels   32;  Gaps    2;

Qy          1 MAAGPRPGAPC-------------------------RAGAPT------IVLTSGRRQTL 28
              : || ||| |                           | |  |      :  |:|| |
Db         51 LVAGSRPGGPIIGYLNLSPPRGAGGAMAELVVHPQSRRRGIGTAMARAALAKTAGRNQFW 110

Qy         29 SHGSSSPARAT 39
              :||:  |||||
Db        111 AHGTLDPARAT 121


RESULT 8
O53831_MYCTU
ID   O53831_MYCTU PRELIMINARY;      PRT;   315 AA.
AC   O53831; Q7D982;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=MT0841, Rv0819;
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium tuberculosis complex.
OX   NCBI_TaxID=1773;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA   Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA   Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA   Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA   Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA   Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA   Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CDC 1551 / Oshkosh;
RX   MEDLINE=22206494; PubMed=12218036;
RX   DOI=10.1128/JB.184.19.5479-5490.2002;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA   Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA   Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA   Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA   Fraser C.M.;
RT   "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   J. Bacteriol. 184:5479-5490(2002).
DR   EMBL; BX842574; CAA17625.1; -; Genomic_DNA.
DR   EMBL; AE000516; AAK45083.1; -; Genomic_DNA.
DR   PIR; C70810; C70810.
DR   PDB; 1OZP; X-ray; A=1-315.
DR   PDB; 1P0H; X-ray; A=1-315.
DR   TIGR; MT0841; -.
DR   TubercuList; Rv0819; -.
DR   GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR   InterPro; IPR000182; GCN5acetyl_trans.
DR   Pfam; PF00583; Acetyltransf_1; 2.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   315 AA;  33599 MW;  E1728C0652CF2716 CRC64;

  Query Match             25.4%;  Score 68;  DB 2;  Length 315;
  Best Local Similarity   28.2%;  Pred. No. 12;
  Matches   20;  Conservative    5;  Mismatches   14;  Indels   32;  Gaps    2;

Qy          1 MAAGPRPGAPC-------------------------RAGAPT------IVLTSGRRQTL 28
              : || ||| |                           | |  |      :  |:|| |
Db         51 LVAGSRPGGPIIGYLNLSPPRGAGGAMAELVVHPQSRRRGIGTAMARAALAKTAGRNQFW 110

Qy         29 SHGSSSPARAT 39
              :||:  |||||
Db        111 AHGTLDPARAT 121


RESULT 9
Q98A00_RHILO
ID   Q98A00_RHILO PRELIMINARY;      PRT;   391 AA.
AC   Q98A00;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Probable dipeptidase.
GN   OrderedLocusNames=mlr6218;
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; BA000012; BAB52544.1; -; Genomic_DNA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0008235; F:metalloexopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000994; Peptidase_M24.
DR   Pfam; PF00557; Peptidase_M24; 1.
KW   Complete proteome; Hydrolase.
SQ   SEQUENCE   391 AA;  43127 MW;  6298B36D3B34A42A CRC64;

  Query Match             25.2%;  Score 67.5;  DB 2;  Length 391;
  Best Local Similarity   41.0%;  Pred. No. 18;
  Matches   16;  Conservative    5;  Mismatches   15;  Indels    3;  Gaps    1;

Qy          6 RPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPL 44
              : ||  |:|    ::||| |  | || ||  |   | |:
```

```
Db        211 KAGADMRSGG---IITSGSRTNLIHGGSSNRRLERGDPV 246


RESULT 10
Q4UVP4_XANCP
ID   Q4UVP4_XANCP PRELIMINARY;      PRT;   563 AA.
AC   Q4UVP4;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   N-acetylmuramoyl-L-alanine amidase.
GN   ORFNames=XC_1816;
OS   Xanthomonas campestris pv. campestris str. 8004.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=314565;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=8004;
RA   Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,
RA   Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
RA   Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,
RA   Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
RA   He C.-Z.;
RL   Submitted (FEB-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CP000050; AAY48879.1; -; Genomic_DNA.
SQ   SEQUENCE   563 AA;  57271 MW;  5860B59189C35941 CRC64;

  Query Match             25.0%;  Score 67;  DB 2;  Length 563;
  Best Local Similarity   46.9%;  Pred. No. 31;
  Matches   15;  Conservative    3;  Mismatches   12;  Indels    2;  Gaps    1;

Qy          5 PRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPA 36
              ||||     |||    | | | |:: |: |||
Db        229 PRPGTVAAPGAP--ATTGGTRTTIATGTPSPA 258


RESULT 11
Q8P8E3_XANCP
ID   Q8P8E3_XANCP PRELIMINARY;      PRT;   563 AA.
AC   Q8P8E3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   N-acetylmuramoyl-L-alanine amidase.
GN   Name=amiC;
OS   Xanthomonas campestris (pv. campestris).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=340;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 33913 / NCPPB 528;
RX   MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA   Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA   Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA   Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA   El-Dorry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA   Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
DR   EMBL; AE012337; AAM41578.1; -; Genomic_DNA.
DR   GO; GO:0008745; F:N-acetylmuramoyl-L-alanine amidase activity; IEA.
DR   GO; GO:0009253; P:peptidoglycan catabolism; IEA.
DR   InterPro; IPR002508; Amidase_3_hydro.
DR   Pfam; PF01520; Amidase_3; 1.
DR   SMART; SM00646; Ami_3; 1.
KW   Complete proteome.
SQ   SEQUENCE   563 AA;  57270 MW;  5860B59189C35941 CRC64;

  Query Match             25.0%;  Score 67;  DB 2;  Length 563;
  Best Local Similarity   46.9%;  Pred. No. 31;
  Matches   15;  Conservative    3;  Mismatches   12;  Indels    2;  Gaps    1;

Qy          5 PRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPA 36
              ||||     |||    | | | |:: |: |||
Db        229 PRPGTVAAPGAP--ATTGGTRTTIATGTPSPA 258


RESULT 12
Q9M4X9_CHLRE
ID   Q9M4X9_CHLRE PRELIMINARY;      PRT;  1787 AA.
AC   Q9M4X9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Flagellar autotomy protein Fa1p.
GN   Name=FA1;
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC   Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20267878; PubMed=10806107;
RA   Finst R.J., Kim P.J., Griffis E.R., Quarmby L.M.;
RT   "Fa1p is a 171 kDa protein essential for axonemal microtubule severing
RT   in Chlamydomonas.";
RL   J. Cell Sci. 113:1963-1971(2000).
DR   EMBL; AF246990; AAF66419.1; -; mRNA.
DR   HSSP; P37595; 1JN9.
DR   GO; GO:0019861; C:flagellum; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR002197; HTH_Fis.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   Pfam; PF00560; LRR_1; 1.
DR   PRINTS; PR01590; HTHFIS.
KW   Flagellum; Leucine-rich repeat; Repeat.
SQ   SEQUENCE   1787 AA;  171585 MW;  452A74259EE14DC1 CRC64;

  Query Match             24.8%;  Score 66.5;  DB 2;  Length 1787;
  Best Local Similarity   38.6%;  Pred. No. 1.3e+02;
  Matches   22;  Conservative    4;  Mismatches   14;  Indels   17;  Gaps    3;

Qy          2 AAG----PRPGAPCRAGAPTIVL-------TSGRRQTL------SHGSSSPARATLG 41
              |||    ||||   :|| | ||          ||:|:|      :  || |   |||
Db        230 AAGGGSAPRPGTAGQAGPATAVLGGSGGLMPPGRQQSLPLSGPSASSSSQPGAVTLG 286


RESULT 13
Q743X3_MYCPA
ID   Q743X3_MYCPA PRELIMINARY;      PRT;   272 AA.
AC   Q743X3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=MAP0468;
OS   Mycobacterium paratuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium avium complex (MAC).
OX   NCBI_TaxID=1770;
```

```
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=k10;
RA   Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE017228; AAS02785.1; -; Genomic_DNA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   InterPro; IPR000073; A/b_hydrolase.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00561; Abhydrolase_1; 1.
KW   Complete proteome; Hydrolase.
SQ   SEQUENCE   272 AA;  30094 MW;  07DB5A382CC8B749 CRC64;

  Query Match             24.6%;  Score 66;  DB 2;  Length 272;
  Best Local Similarity   40.4%;  Pred. No. 18;
  Matches   19;  Conservative    2;  Mismatches    8;  Indels   18;  Gaps    3;

Qy          4 GPRPGAPCR----AGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVL 46
              ||||| | |     | |: |||         ||          |:|||
Db          7 GPRPGQPARCTQDGGMPSTVLT-------RHGGH-------GEPLVL 39


RESULT 14
Q8IV50_HUMAN
ID   Q8IV50_HUMAN PRELIMINARY;      PRT;   215 AA.
AC   Q8IV50;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein MGC35274 (Putative peptidoglycan binding domain
DE   protein-like protein).
GN   Name=MGC35274;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Ding P., Han W., Wang Y., Rui M., Chen Y., Wang L., Ma D.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC033515; AAH33515.1; -; mRNA.
DR   EMBL; AY125955; AAM94507.1; -; mRNA.
DR   Ensembl; ENSG00000140280; Homo sapiens.
DR   GO; GO:0016998; P:cell wall catabolism; IEA.
DR   InterPro; IPR002482; LysM.
DR   Pfam; PF01476; LysM; 1.
DR   SMART; SM00257; LysM; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   215 AA;  23463 MW;  A87E666478E903AC CRC64;

  Query Match             24.1%;  Score 64.5;  DB 2;  Length 215;
  Best Local Similarity   34.0%;  Pred. No. 21;
  Matches   17;  Conservative   11;  Mismatches   13;  Indels    9;  Gaps    2;

Qy          5 PRPGAP-----CRAGAPT----IVLTSGRRQTLSHGSSSPARATLGKPLV 45
              ||| ||       |:|: :    : |:  | :| |:||::  || ||  ::
Db         18 PRPSAPSPPPRSRSGSESEEAELSLSLARTKTRSYGSTASVRAPLGAGVI 67


RESULT 15
Q82NB1_STRAW
ID   Q82NB1_STRAW PRELIMINARY;      PRT;   596 AA.
AC   Q82NB1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative transcriptional regulator.
GN   OrderedLocusNames=SAV1392;
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industrial
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR   EMBL; BA000030; BAC69102.1; -; Genomic_DNA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR   InterPro; IPR003018; GAF.
DR   InterPro; IPR012287; Homeodomain-rel.
DR   InterPro; IPR002197; HTH_Fis.
DR   InterPro; IPR002078; Sig54_interact.
DR   Pfam; PF01590; GAF; 1.
DR   Pfam; PF02954; HTH_8; 1.
DR   Pfam; PF00158; Sigma54_activat; 1.
DR   PRINTS; PR01590; HTHFIS.
DR   TIGRFAMs; TIGR01199; HTH_fis; 1.
DR   PROSITE; PS50045; SIGMA54_INTERACT_4; 1.
KW   Complete proteome.
SQ   SEQUENCE   596 AA;  65277 MW;  ADEE4EAAE22DF77D CRC64;

  Query Match             24.1%;  Score 64.5;  DB 2;  Length 596;
  Best Local Similarity   50.0%;  Pred. No. 66;
  Matches   18;  Conservative    3;  Mismatches   10;  Indels    5;  Gaps    1;

Qy         11 CR-----AGAPTIVLTSGRRQTLSHGSSSPARATLG 41
              ||     |||   |   || |  | ||:||||::||
Db        288 CRREDSDAGAVFRVRLIGRLQPTSAGSASPARSSLG 323
```

Search completed: April 27, 2006, 15:19:40
Job time : 227.824 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 27, 2006, 15:09:06 ; Search time 67.7647 Seconds
(without alignments)
103.742 Million cell updates/sec

Title: US-09-819-144A-8
Perfect score: 83
Sequence: 1 RRQTLSHGSSSPARAC 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	83	100.0	16	2	AAY39295	Aay39295 Polypepti
2	74	89.2	52	2	AAY39293	Aay39293 CSA-1 cho
3	47	56.6	275	3	AAY74422	Aay74422 Neisseria
4	47	56.6	276	3	AAY74421	Aay74421 Neisseria
5	47	56.6	276	3	AAY74420	Aay74420 Neisseria
6	45	54.2	407	6	ABU08934	Abu08934 Human tum
7	45	54.2	407	7	ADM05759	Adm05759 Human pro
8	45	54.2	407	8	ADI79418	Adi79418 Human MAG
9	44	53.0	145	6	AAE37292	Aae37292 Human gen
10	43	51.8	199	7	ABO71484	Abo71484 Pseudomon
11	43	51.8	220	3	AAG32041	Aag32041 Arabidops
12	43	51.8	236	3	AAG32040	Aag32040 Arabidops
13	43	51.8	284	3	AAG32039	Aag32039 Arabidops
14	43	51.8	427	7	ABO73903	Abo73903 Pseudomon
15	43	51.8	854	7	AAE38184	Aae38184 Fruit fly
16	42	50.6	64	4	AAU58002	Aau58002 Propionib
17	42	50.6	64	6	ABM54521	Abm54521 Propionib
18	42	50.6	85	4	AAU21041	Aau21041 Human nov
19	42	50.6	147	7	ABO77336	Abo77336 Pseudomon
20	42	50.6	152	4	AAU40386	Aau40386 Propionib
21	42	50.6	152	6	ABM36905	Abm36905 Propionib
22	42	50.6	215	7	ABO75620	Abo75620 Pseudomon
23	42	50.6	566	8	ADX78475	Adx78475 Plant ful
24	42	50.6	843	7	ABO70181	Abo70181 Pseudomon
25	41.5	50.0	314	5	AAU99417	Aau99417 Human ECS
26	41.5	50.0	345	7	ABO68244	Abo68244 Pseudomon
27	41	49.4	57	4	AAU64156	Aau64156 Propionib
28	41	49.4	57	6	ABM60675	Abm60675 Propionib
29	41	49.4	70	4	AAU51187	Aau51187 Propionib
30	41	49.4	70	6	ABM47706	Abm47706 Propionib
31	41	49.4	112	4	AAU66094	Aau66094 Propionib
32	41	49.4	112	6	ABM62613	Abm62613 Propionib
33	41	49.4	155	5	ABP55326	Abp55326 Human qui
34	41	49.4	178	5	ABP41446	Abp41446 Human ova
35	41	49.4	213	7	ABO73901	Abo73901 Pseudomon
36	41	49.4	349	7	ABO81915	Abo81915 Pseudomon
37	41	49.4	412	8	ADT60756	Adt60756 Plant pol
38	41	49.4	470	5	AAE19560	Aae19560 Equine he
39	41	49.4	740	4	AAB94340	Aab94340 Human pro
40	41	49.4	740	7	ADJ70657	Adj70657 Human hea
41	41	49.4	813	3	AAG52554	Aag52554 Arabidops
42	41	49.4	831	7	ABO76155	Abo76155 Pseudomon
43	41	49.4	835	4	AAB35401	Aab35401 Human PG-
44	41	49.4	835	5	ABP69678	Abp69678 Human pol
45	41	49.4	835	5	ABP53317	Abp53317 Human PG-

ALIGNMENTS

RESULT 1
AAY39295
ID AAY39295 standard; peptide; 16 AA.
XX
AC AAY39295;
XX
DT 26-NOV-1999 (first entry)
XX
DE Polypeptide used to generate CSA-1-reactive polyclonal antibodies.
XX
KW Chondrosarcoma associated protein-1; CSA-1; bone malignancy; diagnosis;
KW rheumatoid arthritis; inflammatory arthropathy; tumour; immunogen;
KW cartilage associated polypeptide; CAA-1; polyclonal antibody generation.
XX
OS Homo sapiens.
XX
PN WO9946382-A1.
XX
PD 16-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US005348.
XX
PR 13-MAR-1998; 98US-00042225.
XX
PA (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.
XX
PI Terek RM;
XX
DR WPI; 1999-551411/46.
XX
PT New gene encoding a chondrosarcoma associated (CSA) polypeptide, useful
PT for diagnosing bone malignancy.
XX
PS Example 3; Page 24; 47pp; English.
XX
CC This sequence is a peptide used as an immunogen to stimulate the
CC production of human chondrosarcoma associated protein-1 (CSA-1:AAY39293)
CC reactive polyclonal antibodies. Chondrosarcoma is the second most common
CC form of bone malignancy and occurs in late adulthood and old age. CSA-1
CC is expressed in a tumour cell line and also in some high grade
CC chondrosarcoma, but not in normal cartilage, or low or intermediate grade
CC tumours. The CSA-1 polynucleotide and polypeptide can be used in methods
CC and compositions for evaluating appropriate treatment and treatment
CC effectiveness of malignancies associated with expression of CSA-1. CSA-1
CC polynucleotide can be used as a probe to classify cells in terms of their
CC level of CSA-1 expression or as primers for diagnostic PCR analysis in
CC which mutations and allelic variation of CSA-1 can be detected.

```
CC   Transgenic animals containing human CSA-1 or with a null mutation can
CC   serve as models for chondrosarcoma. Methods of treating undesired
CC   inflammation such as that associated with rheumatoid arthritis and other
CC   inflammatory arthopathies is carried out by administering an effective
CC   amount of a cartilage associated (CAA) polypeptide such as CAA-1 AAY39294
XX
SQ   Sequence 16 AA;

  Query Match             100.0%;  Score 83;  DB 2;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARAC 16
              ||||||||||||||||
Db          1 RRQTLSHGSSSPARAC 16


RESULT 2
AAY39293
ID   AAY39293 standard; protein; 52 AA.
XX
AC   AAY39293;
XX
DT   26-NOV-1999  (first entry)
XX
DE   CSA-1 chondrosarcoma associated protein-1.
XX
KW   Chondrosarcoma associated protein-1; CSA-1; bone malignancy; diagnosis;
KW   rheumatoid arthritis; inflammatory arthropathy; tumour;
KW   cartilage associated polypeptide; CAA-1.
XX
OS   Homo sapiens.
XX
PN   WO9946382-A1.
XX
PD   16-SEP-1999.
XX
PF   12-MAR-1999;   99WO-US005348.
XX
PR   13-MAR-1998;   98US-00042225.
XX
PA   (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.
XX
PI   Terek RM;
XX
DR   WPI; 1999-551411/46.
DR   N-PSDB; AAZ06793.
XX
PT   New gene encoding a chondrosarcoma associated (CSA) polypeptide, useful
PT   for diagnosing bone malignancy.
XX
PS   Claim 10; Page 18; 47pp; English.
XX
CC   This is the human chondrosarcoma associated protein-1 (CSA-1) amino acid
CC   sequence. Chondrosarcoma is the second most common form of bone
CC   malignancy and occurs in late adulthood and old age. CSA-1 is expressed
CC   in a tumour cell line and also in some high grade chondrosarcoma, but not
CC   in normal cartilage, or low or intermediate grade tumours. The CSA-1
CC   polynucleotide and polypeptide can be used in methods and compositions
CC   for evaluating appropriate treatment and treatment effectiveness of
CC   malignancies associated with expression of CSA-1. CSA-1 polynucleotide
CC   can be used as a probe to classify cells in terms of their level of CSA-1
CC   expression or as primers for diagnostic PCR analysis in which mutations
CC   and allelic variation of CSA-1 can be detected. Transgenic animals
CC   containing human CSA-1 or with a null mutation can serve as models for
CC   chondrosarcoma. Methods of treating undesired inflammation such as that
CC   associated with rheumatoid arthritis and other inflammatory arthopathies
CC   is carried out by administering an effective amount of a cartilage
CC   associated (CAA) polypeptide such as CAA-1 AAY39294
XX
SQ   Sequence 52 AA;

  Query Match             89.2%;  Score 74;  DB 2;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARA 15
              |||||||||||||||
Db         24 RRQTLSHGSSSPARA 38


RESULT 3
AAY74422
ID   AAY74422 standard; protein; 275 AA.
XX
AC   AAY74422;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 093 protein sequence SEQ ID NO:320.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US009346.
XX
PR   01-MAY-1998;   98US-0083758P.
PR   31-JUL-1998;   98US-0094869P.
PR   02-SEP-1998;   98US-0098994P.
PR   02-SEP-1998;   98US-0099062P.
PR   09-OCT-1998;   98US-0103749P.
PR   09-OCT-1998;   98US-0103794P.
PR   09-OCT-1998;   98US-0103796P.
PR   25-FEB-1999;   99US-0121528P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   N-PSDB; AAZ53184.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics.
XX
PS   Claim 2; Page 294; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of the
CC   invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the manufacture
CC   of medicaments for treating or preventing infection due to Neisserial
CC   bacteria (e.g. meningitis and septicaemia), to detect the presence of
CC   Neisseria bacteria, or to raise antibodies. They may also be used to
CC   screen for agonists or antagonists, which may themselves have use as
CC   antibacterial agents. The polynucleotides of the invention may also be
CC   used in gene therapy protocols
XX
SQ   Sequence 275 AA;

  Query Match             56.6%;  Score 47;  DB 3;  Length 275;
  Best Local Similarity   71.4%;  Pred. No. 12;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy          1 RRQTLSHGSSSPAR 14
              |||||| |:  |||
Db        258 RRQTLSVGNQHPAR 271
```

RESULT 4
AAY74421
ID AAY74421 standard; protein; 276 AA.
XX
AC AAY74421;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 093 protein sequence SEQ ID NO:318.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US009346.
XX
PR 01-MAY-1998; 98US-0083758P.
PR 31-JUL-1998; 98US-0094869P.
PR 02-SEP-1998; 98US-0098994P.
PR 02-SEP-1998; 98US-0099062P.
PR 09-OCT-1998; 98US-0103749P.
PR 09-OCT-1998; 98US-0103794P.
PR 09-OCT-1998; 98US-0103796P.
PR 25-FEB-1999; 99US-0121528P.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR N-PSDB; AAZ53183.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics.
XX
PS Claim 2; Page 293; 1453pp; English.
XX
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of the
CC invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the manufacture
CC of medicaments for treating or preventing infection due to Neisserial
CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
CC Neisseria bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols
XX
SQ Sequence 276 AA;

```
  Query Match             56.6%;  Score 47;  DB 3;  Length 276;
  Best Local Similarity   71.4%;  Pred. No. 12;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPAR 14
              |||||| |:  |||
Db        259 RRQTLSVGNQHPAR 272
```

RESULT 5
AAY74420
ID AAY74420 standard; protein; 276 AA.
XX
AC AAY74420;
XX
DT 12-SEP-2003 (revised)
DT 21-MAR-2000 (first entry)
XX
DE Neisseria gonorrheae ORF 093 protein sequence SEQ ID NO:316.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US009346.
XX
PR 01-MAY-1998; 98US-0083758P.
PR 31-JUL-1998; 98US-0094869P.
PR 02-SEP-1998; 98US-0098994P.
PR 02-SEP-1998; 98US-0099062P.
PR 09-OCT-1998; 98US-0103749P.
PR 09-OCT-1998; 98US-0103794P.
PR 09-OCT-1998; 98US-0103796P.
PR 25-FEB-1999; 99US-0121528P.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR N-PSDB; AAZ53182.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics.
XX
PS Claim 2; Page 292; 1453pp; English.
XX
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of the
CC invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the manufacture
CC of medicaments for treating or preventing infection due to Neisserial
CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
CC Neisseria bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
CC field)
XX
SQ Sequence 276 AA;

```
  Query Match             56.6%;  Score 47;  DB 3;  Length 276;
  Best Local Similarity   71.4%;  Pred. No. 12;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPAR 14
```

```
                ||||||  |:  |||
Db        259 RRQTLSVGNQHPAR 272
```

```
RESULT 6
ABU08934
ID   ABU08934 standard; protein; 407 AA.
XX
AC   ABU08934;
XX
DT   05-JUN-2003  (first entry)
XX
DE   Human tumour rejection antigen precursor, MAGE-B6.
XX
KW   TRAP; tumour rejection antigen precursor; cytolytic T-cell; CTL; tumour;
KW   seminoma; bladder transitional-cell carcinoma; NSCLC; adaptor;
KW   head-and-neck squamous-cell carcinoma; breast carcinoma; sarcoma;
KW   cutaneous melanoma; nonsmall cell lung cancer; MAGE-B6; human.
XX
OS   Homo sapiens.
XX
PN   US2002176865-A1.
XX
PD   28-NOV-2002.
XX
PF   01-MAR-2002; 2002US-00085108.
XX
PR   25-APR-1997;   97US-00845528.
PR   24-APR-1998;   98US-00066281.
PR   17-DEC-1999;   99US-00468433.
PR   09-FEB-2000; 2000US-00501104.
XX
PA   (LUCA/) LUCAS S.
PA   (BOON/) BOON-FALLEUR T.
XX
PI   Lucas S,  Boon-Falleur T;
XX
DR   WPI; 2003-328468/31.
DR   N-PSDB; ABX95008.
XX
PT   Novel isolated nucleic acid encoding tumor rejection antigen precursor
PT   MAGE-C3, MAGE-B5, or MAGE-B6, useful as diagnostic probes to determine
PT   presence of abnormal e.g., tumor cells expressing MAGE-C1, MAGE-B5 or
PT   MAGE-B6.
XX
PS   Example 13; Fig 10; 59pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule which encodes
CC   a tumour rejection antigen precursor (TRAP) having an amino acid sequence
CC   of a TRAP encoded by a fully defined MAGE-C3, MAGE-B5, or MAGE-B6
CC   polynucleotide sequence. Also disclosed is a method which is useful for
CC   determining presence of cytolytic T-cells specific for complexes of human
CC   leukocyte antigen (HLA) and a peptide derived from the nucleic acid in a
CC   cytotoxic T-lymphocyte (CTL)-containing sample. The nucleic acid is
CC   useful as a diagnostic probe to determine the presence of abnormal
CC   (tumour) cells such as seminoma, bladder transitional-cell carcinoma,
CC   head-and-neck squamous-cell carcinoma, breast carcinoma, sarcoma,
CC   cutaneous melanoma or nonsmall cell lung cancer (NSCLC) which express
CC   MAGE-C1, MAGE-B5 or MAGE-B6. The nucleic acid is useful for diagnosing a
CC   disorder characterised by expression of MAGE-C1, MAGE-B5 or MAGE-B6 TRAPs
CC   or tumour rejection antigens (TRAs). The present sequence represents the
CC   amino acid sequence of the human tumour rejection antigen precursor, MAGE
CC   -B6
XX
SQ   Sequence 407 AA;
```

```
  Query Match             54.2%;  Score 45;  DB 6;  Length 407;
  Best Local Similarity   60.0%;  Pred. No. 42;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              :|  || ||| :|||
Db         34 KQEESHSSSSSSRAC 48
```

```
RESULT 7
ADM05759
ID   ADM05759 standard; protein; 407 AA.
XX
AC   ADM05759;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human protein of the invention SEQ ID NO:4444.
XX
KW   human; gene therapy; diagnostic marker; pharmaceutical.
XX
OS   Homo sapiens.
XX
PN   EP1347046-A1.
XX
PD   24-SEP-2003.
XX
PF   12-APR-2002; 2002EP-00008400.
XX
PR   22-MAR-2002; 2002JP-00137785.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Sugiyama T,  Otsuki T,  Wakamatsu A,  Sato H,  Ishii S;
PI   Yamamoto J,  Isono Y,  Hio Y,  Otsuka K,  Nagai K,  Irie R,  Tamechika I;
PI   Seki N,  Yoshikawa T,  Otsuka M,  Nagahari K,  Masuho Y;
XX
DR   WPI; 2003-723558/69.
DR   N-PSDB; ADM03316.
XX
PT   New polynucleotides and polypeptides are useful in gene therapy, for
PT   developing a diagnostic marker or medicines for regulating their
PT   expression and activity, or as a target of gene therapy.
XX
PS   Claim 1; SEQ ID NO 4444; 305pp; English.
XX
CC   The invention relates to a novel human polynucleotide and the encoded
CC   polypeptide. A polynucleotide of the invention may have a use in gene
CC   therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC   as a primer for synthesizing the polynucleotide or as a probe for
CC   detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC   useful in gene therapy, for developing a diagnostic marker or medicines
CC   for regulating their expression and activity, or as a target of gene
CC   therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC   are useful as pharmaceutical agents. The present sequence represents a
CC   protein sequence of the invention.
XX
SQ   Sequence 407 AA;
```

```
  Query Match             54.2%;  Score 45;  DB 7;  Length 407;
  Best Local Similarity   60.0%;  Pred. No. 42;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              :|  || ||| :|||
Db         34 KQEESHSSSSSSRAC 48
```

```
RESULT 8
ADI79418
ID   ADI79418 standard; protein; 407 AA.
XX
AC   ADI79418;
XX
DT   22-APR-2004  (first entry)
XX
DE   Human MAGE-B6 protein.
XX
KW   Human; MAGE-B6; cancer; cytostatic; TRAP;
KW   tumour rejection antigen precursor.
```

XX
OS Homo sapiens.
XX
PN US6680191-B1.
XX
PD 20-JAN-2004.
XX
PF 17-DEC-1999; 99US-00468433.
XX
PR 25-APR-1997; 97US-00845528.
PR 24-APR-1998; 98US-00066281.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Lucas S, Boon-Falleur T;
XX
DR WPI; 2004-088565/09.
DR N-PSDB; ADI79417.
XX
PT New nucleic acid molecules coding for tumor rejection antigen precursors
PT of the MAGE-C and MAGE-B families, useful for diagnosing, preventing or
PT treating cancer.
XX
PS Example 13; SEQ ID NO 26; 56pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprising the
CC open reading frame of human MAGE-C3 (not defined) appearing as ADI79413,
CC or its complete complement. Also included are an expression vector
CC comprising the new nucleic acid molecule operatively linked to a
CC promoter, an isolated cell line or cell strain transfected or transformed
CC with the expression vector and a kit useful in a polymerase chain
CC reaction (PCR) based assay, comprising an oligonucleotide fragment of
CC ADI79413 comprising nucleotides 175-195 or 711-731. MAGE-C and MAGE-B
CC family members are tumour rejection antigen precursors (TRAP). The
CC composition and methods are useful for diagnosing, preventing or treating
CC cancer. Also disclosed as new are the DNAs and proteins for MAGE-C1, MAGE
CC -C2, MAGE-B5 and MAGE-B6. The genes for MAGE-C1, C2 and C3 are located on
CC chromosome Xq26-q27. The present sequence represents MAGE-B6.
XX
SQ Sequence 407 AA;

Query Match 54.2%; Score 45; DB 8; Length 407;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy          2 RQTLSHGSSSPARAC 16
              :|  || ||| :|||
Db         34 KQEESHSSSSSSRAC 48
```

RESULT 9
AAE37292
ID AAE37292 standard; protein; 145 AA.
XX
AC AAE37292;
XX
DT 07-AUG-2003 (first entry)
XX
DE Human gene 9 encoded secreted protein HDMSQ09, SEQ ID NO:55.
XX
KW Human; secreted protein; hyperproliferative disorder; multiple myeloma;
KW immune disorder; cancer; acquired immune deficiency syndrome; arthritis;
KW haematopoietic disorder; systemic lupus erythematosus; atherosclerosis;
KW reproductive system disorder; musculoskeletal disorder; renal disorder;
KW cardiovascular disorder; neurofibromatosis; Alzheimer's disease; asthma;
KW neurological disorder; Parkinson's disease; respiratory disorder; AIDS;
KW emphysema; bronchitis; endocrine disorder; genetic disorder; psoriasis;
KW inflammation; congestive heart failure; multiple sclerosis; arrhythmia;
KW hypertension; urolithiasis; amenorrhoea; epididymitis; anaemia; allergy;
KW gout; hepatitis; digestive disease; diabetes; goitre; infection; wound;
KW gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1
FT /label= Signal_peptide
FT Protein 2. .145
FT /note= "Mature human secreted protein"
XX
PN WO2003038038-A2.
XX
PD 08-MAY-2003.
XX
PF 24-OCT-2002; 2002WO-US033985.
XX
PR 26-OCT-2001; 2001US-0330629P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Birse CE, Komatsoulis GA, Choi GH;
XX
DR WPI; 2003-430509/40.
DR N-PSDB; AAD56371.
XX
PT New isolated nucleic acid molecule and human secreted proteins, useful
PT for diagnosing, preventing or treating diseases e.g. cancer,
PT atherosclerosis, anemia, multiple myeloma, hypertension, infections or
PT wounds.
XX
PS Claim 11; Page 461; 468pp; English.
XX
CC AAD56351-AAD56374 represent cDNAs corresponding to 16 human secreted
CC protein genes and AAE37272-AAE37295 represent the proteins they encode.
CC AAE37296-AAE37299 represent human secreted protein fragments. The genes
CC and their corresponding proteins are useful in diagnosing, preventing or
CC treating diseases or conditions such as hyperproliferative disorders
CC (e.g. cancer), immune/haematopoietic disorders (e.g. anaemia, multiple
CC myeloma, arthritis, asthma, acquired immune deficiency syndrome (AIDS),
CC rheumatoid arthritis, inflammation, psoriasis, diabetes, systemic lupus
CC erythematosus or allergy), reproductive system disorders (e.g.
CC amenorrhoea or epididymitis), musculoskeletal disorders (e.g. gout or
CC multiple sclerosis), cardiovascular disorders (e.g. arrhythmia,
CC congestive heart failure, hypertension or atherosclerosis), genetic
CC disorders (e.g. neurofibromatosis), renal disorders (e.g. urolithiasis),
CC neurological disorders (e.g. Alzheimer's disease or Parkinson's disease),
CC respiratory disorders (e.g. emphysema or bronchitis), endocrine disorders
CC (e.g. goitre or diabetes), digestive diseases (e.g. hepatitis),
CC infections or wounds. Polynucleotides of the invention may also be used
CC for chromosome mapping or drug screening. They are also useful in gene
CC therapy. The present sequence represents a human secreted protein of the
CC invention
XX
SQ Sequence 145 AA;

Query Match 53.0%; Score 44; DB 6; Length 145;
Best Local Similarity 56.2%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 RRQTLSHGSSSPARAC 16
                |  |||| | |  ||
Db         36 RFPVLSHGLSGPPHAC 51
```

RESULT 10
ABO71484
ID ABO71484 standard; protein; 199 AA.
XX
AC ABO71484;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #3659.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX

OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD05055.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 20230; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 199 AA;

Query Match 51.8%; Score 43; DB 7; Length 199;
Best Local Similarity 53.3%; Pred. No. 44;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```
Qy          2 RQTLSHGSSSPARAC 16
              | |:: ||| | | |
Db        150 RWTIASGSSPPCRRC 164
```

RESULT 11
AAG32041
ID AAG32041 standard; protein; 220 AA.
XX
AC AAG32041;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38581.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 51.8%; Score 43; DB 3; Length 220;
Best Local Similarity 53.3%; Pred. No. 49;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 RRQTLSHGSSSPARA 15
              :|:| |||: ||  |
Db         39 KRKTCSHGTRSPVLA 53
```

RESULT 12
AAG32040
ID AAG32040 standard; protein; 236 AA.
XX
AC AAG32040;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38580.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.

PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 51.8%; Score 43; DB 3; Length 236;
Best Local Similarity 53.3%; Pred. No. 52;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 RRQTLSHGSSSPARA 15
              :|:| |||: ||  |
Db         55 KRKTCSHGTRSPVLA 69
```

RESULT 13
AAG32039
ID AAG32039 standard; protein; 284 AA.
XX
AC AAG32039;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38579.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 51.8%; Score 43; DB 3; Length 284;
Best Local Similarity 53.3%; Pred. No. 64;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 RRQTLSHGSSSPARA 15
              :|:| |||: ||  |
Db        103 KRKTCSHGTRSPVLA 117
```

RESULT 14
ABO73903
ID ABO73903 standard; protein; 427 AA.
XX
AC ABO73903;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #6078.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD07474.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 22649; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of

```
CC   effective antibacterial targets, as targets for antibacterial drugs,
CC   including anti-P. aeruginosa drugs, as templates for recombinant
CC   production of P. aeruginosa-derived peptides or polypeptides, as target
CC   components for diagnosis and/or treatment of P. aeruginosa-caused
CC   infection, and in detection of P. aeruginosa sequences or other sequences
CC   of Pseudomonas species using biochip technology. Sequences ABO67826-
CC   ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format from USPTO at
CC   seqdata.uspto.gov/sequence.html
XX
SQ   Sequence 427 AA;

  Query Match             51.8%;  Score 43;  DB 7;  Length 427;
  Best Local Similarity   53.3%;  Pred. No. 98;
  Matches    8;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              ||   |||  ||: |
Db        219 RQGQGHGSGHPAQPC 233


RESULT 15
AAE38184
ID   AAE38184 standard; protein; 854 AA.
XX
AC   AAE38184;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Fruit fly G protein-coupled receptor (GPCR) protein #33.
XX
KW   Fruit fly; G protein-coupled receptor; GPCR; cell therapy; nematode;
KW   insect infestation; insecticide; nematocide; receptor.
XX
OS   Drosophila melanogaster.
XX
PN   WO2003052078-A2.
XX
PD   26-JUN-2003.
XX
PF   18-DEC-2002; 2002WO-US040525.
XX
PR   18-DEC-2001; 2001US-0341512P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Spana E,  Kamdar K,  Stam L,  Valentine S,  Griswold CM;
XX
DR   WPI; 2003-533019/50.
DR   N-PSDB; AAD57461.
XX
PT   New polypeptide, useful for preventing or abrogating insect or nematode
PT   infestation of a plant.
XX
PS   Claim 1; Page 250-254; 368pp; English.
XX
CC   The invention relates to a novel G protein-coupled receptor (GPCR)
CC   polypeptide and its polynucleotide. The polypeptide is useful for
CC   preventing or abrogating insect or nematode infestation of a plant. GPCR
CC   polypeptide is useful in cell therapy. The present sequence is fruit fly
CC   GPCR protein
XX
SQ   Sequence 854 AA;

  Query Match             51.8%;  Score 43;  DB 7;  Length 854;
  Best Local Similarity   66.7%;  Pred. No. 2e+02;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 LSHGSSSPARAC 16
              | | |||||  :|
Db        603 LRHSSSSPASSC 614
```

Search completed: April 27, 2006, 15:14:42
Job time : 69.7647 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 27, 2006, 15:20:00 ; Search time 15.5294 Seconds
(without alignments)
85.181 Million cell updates/sec

Title: US-09-819-144A-8
Perfect score: 83
Sequence: 1 RRQTLSHGSSSPARAC 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	83	100.0	16	2	US-09-042-225-8	Sequence 8, Appli
2	74	89.2	52	2	US-09-042-225-2	Sequence 2, Appli
3	45	54.2	407	2	US-09-468-433C-26	Sequence 26, Appl
4	44	53.0	336	2	US-09-270-767-34524	Sequence 34524, A
5	44	53.0	336	2	US-09-270-767-49741	Sequence 49741, A
6	43	51.8	199	2	US-09-252-991A-20230	Sequence 20230, A
7	43	51.8	339	2	US-09-270-767-45916	Sequence 45916, A
8	43	51.8	427	2	US-09-252-991A-22649	Sequence 22649, A
9	42	50.6	147	2	US-09-252-991A-26082	Sequence 26082, A
10	42	50.6	215	2	US-09-252-991A-24366	Sequence 24366, A
11	42	50.6	843	2	US-09-252-991A-18927	Sequence 18927, A
12	41.5	50.0	345	2	US-09-252-991A-16990	Sequence 16990, A
13	41	49.4	67	2	US-09-621-976-6000	Sequence 6000, Ap
14	41	49.4	213	2	US-09-252-991A-22647	Sequence 22647, A
15	41	49.4	349	2	US-09-252-991A-30661	Sequence 30661, A
16	41	49.4	470	2	US-10-332-795-3	Sequence 3, Appli
17	41	49.4	831	2	US-09-252-991A-24901	Sequence 24901, A
18	40	48.2	163	2	US-09-252-991A-22928	Sequence 22928, A
19	40	48.2	248	2	US-09-252-991A-22391	Sequence 22391, A
20	40	48.2	562	2	US-09-879-792-12	Sequence 12, Appl
21	40	48.2	1054	2	US-09-949-016-9821	Sequence 9821, Ap
22	40	48.2	1054	2	US-09-949-016-9822	Sequence 9822, Ap
23	40	48.2	1262	2	US-09-357-251-33	Sequence 33, Appl
24	40	48.2	1262	2	US-09-949-016-6182	Sequence 6182, Ap
25	40	48.2	1262	2	US-09-949-016-6850	Sequence 6850, Ap
26	40	48.2	1266	1	US-08-468-557-4	Sequence 4, Appli
27	40	48.2	1266	2	US-09-357-251-32	Sequence 32, Appl
28	40	48.2	1621	1	US-08-242-677-2	Sequence 2, Appli
29	39	47.0	100	2	US-09-390-134B-39	Sequence 39, Appl
30	39	47.0	125	2	US-09-621-976-5826	Sequence 5826, Ap
31	39	47.0	134	2	US-09-252-991A-17233	Sequence 17233, A
32	39	47.0	147	2	US-09-252-991A-28936	Sequence 28936, A
33	39	47.0	178	2	US-09-270-767-32556	Sequence 32556, A
34	39	47.0	178	2	US-09-270-767-47773	Sequence 47773, A
35	39	47.0	186	2	US-09-673-763-12	Sequence 12, Appl
36	39	47.0	212	2	US-09-252-991A-30730	Sequence 30730, A
37	39	47.0	251	2	US-09-252-991A-28124	Sequence 28124, A
38	39	47.0	461	2	US-09-422-936-57	Sequence 57, Appl
39	39	47.0	496	2	US-09-422-936-85	Sequence 85, Appl
40	39	47.0	563	2	US-09-248-796A-17571	Sequence 17571, A
41	39	47.0	578	2	US-09-422-936-55	Sequence 55, Appl
42	39	47.0	581	2	US-09-422-936-59	Sequence 59, Appl
43	39	47.0	581	2	US-09-252-991A-20966	Sequence 20966, A
44	39	47.0	776	2	US-09-252-991A-26717	Sequence 26717, A
45	39	47.0	844	2	US-09-422-936-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1
US-09-042-225-8
; Sequence 8, Application US/09042225A
; Patent No. 6207812
; GENERAL INFORMATION:
; APPLICANT: Terek, Richard M.
; TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES
; FILE REFERENCE: 04930/021001
; CURRENT APPLICATION NUMBER: US/09/042,225A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-042-225-8

Query Match 100.0%; Score 83; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.6e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRQTLSHGSSSPARAC 16
||||||||||||||||
Db 1 RRQTLSHGSSSPARAC 16

RESULT 2
US-09-042-225-2
; Sequence 2, Application US/09042225A
; Patent No. 6207812
; GENERAL INFORMATION:
; APPLICANT: Terek, Richard M.
; TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES
; FILE REFERENCE: 04930/021001
; CURRENT APPLICATION NUMBER: US/09/042,225A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-042-225-2

Query Match 89.2%; Score 74; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRQTLSHGSSSPARA 15

```
              ||||||||||||||||
Db         24 RRQTLSHGSSSPARA 38


RESULT 3
US-09-468-433C-26
; Sequence 26, Application US/09468433C
; Patent No. 6680191
;  GENERAL INFORMATION:
;    APPLICANT:  LUCAS, Sophie; BOON-FALLEUR, Thierry
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULES CODING FOR
;    TITLE OF INVENTION:  TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C AN
;    TITLE OF INVENTION:  MAGE-B FAMILIES AND USES THEREOF
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski L.L.P.
;      STREET:  801 Pennsylvania Avenue, NW
;      CITY:  Washington
;      STATE:  District of Columbia
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/468,433C
;      FILING DATE:  December 17, 1999
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/066,281
;      FILING DATE:  April 24, 1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/845,528
;      FILING DATE:  April 25, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary Anne Schofield
;      REGISTRATION NUMBER:  36,669
;      REFERENCE/DOCKET NUMBER:  LUD 5611 JEL/MAS
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 662-0200
;      TELEFAX:  (202) 662-4643
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  407
;      TYPE:  amino acid
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
US-09-468-433C-26

  Query Match             54.2%;  Score 45;  DB 2;  Length 407;
  Best Local Similarity   60.0%;  Pred. No. 13;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              :|  || ||| :|||
Db         34 KQEESHSSSSSSRAC 48


RESULT 4
US-09-270-767-34524
; Sequence 34524, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34524
;   LENGTH: 336
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-34524

  Query Match             53.0%;  Score 44;  DB 2;  Length 336;
  Best Local Similarity   81.8%;  Pred. No. 15;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSS 11
              | |||||| |
Db         22 RSQTLSHGSGS 32


RESULT 5
US-09-270-767-49741
; Sequence 49741, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49741
;   LENGTH: 336
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-49741

  Query Match             53.0%;  Score 44;  DB 2;  Length 336;
  Best Local Similarity   81.8%;  Pred. No. 15;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSS 11
              | |||||| |
Db         22 RSQTLSHGSGS 32


RESULT 6
US-09-252-991A-20230
; Sequence 20230, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 20230
;   LENGTH: 199
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20230

  Query Match             51.8%;  Score 43;  DB 2;  Length 199;
  Best Local Similarity   53.3%;  Pred. No. 12;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              | |:: ||| | | |
Db        150 RWTIASGSSPPCRRC 164


RESULT 7
```

```
US-09-270-767-45916
; Sequence 45916, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45916
;   LENGTH: 339
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
;   FEATURE:
;   OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-45916

  Query Match             51.8%;  Score 43;  DB 2;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 23;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 QTLSHGSSSPAR 14
              ||:|| : ||||
Db        164 QTISHRAQSPAR 175


RESULT 8
US-09-252-991A-22649
; Sequence 22649, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 22649
;   LENGTH: 427
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22649

  Query Match             51.8%;  Score 43;  DB 2;  Length 427;
  Best Local Similarity   53.3%;  Pred. No. 29;
  Matches    8;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              ||   |||  ||: |
Db        219 RQGQGHGSGHPAQPC 233


RESULT 9
US-09-252-991A-26082
; Sequence 26082, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 26082
;   LENGTH: 147
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26082

  Query Match             50.6%;  Score 42;  DB 2;  Length 147;
  Best Local Similarity   60.0%;  Pred. No. 13;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 HGSSSPARAC 16
              ||:::||| |
Db         24 HGATTPARCC 33


RESULT 10
US-09-252-991A-24366
; Sequence 24366, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 24366
;   LENGTH: 215
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24366

  Query Match             50.6%;  Score 42;  DB 2;  Length 215;
  Best Local Similarity   53.3%;  Pred. No. 20;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              |: |  |||:| | |
Db         14 RRPLPRGSSAPRRRC 28


RESULT 11
US-09-252-991A-18927
; Sequence 18927, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 18927
;   LENGTH: 843
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18927

  Query Match             50.6%;  Score 42;  DB 2;  Length 843;
  Best Local Similarity   50.0%;  Pred. No. 96;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy          3 QTLSHGSSSPARAC 16
              ||:||   :|: ||
Db        326 QTVSHAQPTPSGAC 339


RESULT 12
US-09-252-991A-16990
; Sequence 16990, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 16990
;   LENGTH: 345
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16990

  Query Match             50.0%;  Score 41.5;  DB 2;  Length 345;
  Best Local Similarity   58.8%;  Pred. No. 42;
  Matches   10;  Conservative    0;  Mismatches    4;  Indels    3;  Gaps    1;

Qy          1 RRQTLSHGSS---SPAR 14
              ||| ||||       |||
Db        307 RRQALSHGGGQLPDPAR 323


RESULT 13
US-09-621-976-6000
; Sequence 6000, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 6000
;   LENGTH: 67
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-621-976-6000

  Query Match             49.4%;  Score 41;  DB 2;  Length 67;
  Best Local Similarity   53.3%;  Pred. No. 7.8;
  Matches    8;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              |  | :  | |||||
Db         47 RSALGNNHSPPARAC 61


RESULT 14
US-09-252-991A-22647
; Sequence 22647, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 22647
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22647

  Query Match             49.4%;  Score 41;  DB 2;  Length 213;
  Best Local Similarity   50.0%;  Pred. No. 29;
  Matches    8;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARAC 16
              |||  : | |:| | |
Db        109 RRQPRTRGWSAPRRGC 124


RESULT 15
US-09-252-991A-30661
; Sequence 30661, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 30661
;   LENGTH: 349
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30661

  Query Match             49.4%;  Score 41;  DB 2;  Length 349;
  Best Local Similarity   57.1%;  Pred. No. 52;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPAR 14
              :|:|  |||| | |
Db         44 QRRTKDHGSSLPIR 57


Search completed: April 27, 2006, 15:21:44
Job time : 16.5294 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 27, 2006, 15:35:30 ; Search time 56 Seconds
(without alignments)
119.380 Million cell updates/sec

Title: US-09-819-144A-8
Perfect score: 83
Sequence: 1 RRQTLSHGSSSPARAC 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	83	100.0	16	3	US-09-819-144A-8	Sequence 8, Appli
2	74	89.2	52	3	US-09-819-144A-2	Sequence 2, Appli
3	51	61.4	415	4	US-10-437-963-145735	Sequence 145735,
4	45	54.2	362	4	US-10-437-963-145733	Sequence 145733,
5	45	54.2	407	4	US-10-085-108-26	Sequence 26, Appl
6	45	54.2	407	4	US-10-108-260A-4444	Sequence 4444, Ap
7	44	53.0	145	5	US-10-773-236-320	Sequence 320, App
8	44	53.0	168	4	US-10-425-115-213570	Sequence 213570,
9	44	53.0	1563	4	US-10-437-963-140084	Sequence 140084,
10	43	51.8	138	4	US-10-437-963-144158	Sequence 144158,
11	43	51.8	198	4	US-10-425-115-199021	Sequence 199021,
12	43	51.8	378	4	US-10-437-963-143718	Sequence 143718,
13	43	51.8	711	4	US-10-437-963-134466	Sequence 134466,
14	43	51.8	735	4	US-10-437-963-120904	Sequence 120904,
15	43	51.8	854	5	US-10-489-425-66	Sequence 66, Appl
16	42	50.6	47	4	US-10-425-115-358602	Sequence 358602,
17	42	50.6	153	4	US-10-437-963-128233	Sequence 128233,
18	42	50.6	233	4	US-10-425-115-275385	Sequence 275385,
19	42	50.6	566	4	US-10-425-114-47841	Sequence 47841, A
20	41.5	50.0	314	4	US-10-416-090-26	Sequence 26, Appl
21	41	49.4	43	4	US-10-425-115-297323	Sequence 297323,
22	41	49.4	73	4	US-10-767-701-61697	Sequence 61697, A
23	41	49.4	89	4	US-10-437-963-120393	Sequence 120393,
24	41	49.4	89	4	US-10-425-115-357735	Sequence 357735,
25	41	49.4	90	4	US-10-425-115-239337	Sequence 239337,
26	41	49.4	146	4	US-10-437-963-152522	Sequence 152522,
27	41	49.4	157	4	US-10-425-115-359138	Sequence 359138,
28	41	49.4	176	4	US-10-437-963-138329	Sequence 138329,
29	41	49.4	178	4	US-10-264-049-2578	Sequence 2578, Ap
30	41	49.4	189	4	US-10-437-963-149015	Sequence 149015,
31	41	49.4	203	4	US-10-424-599-258595	Sequence 258595,
32	41	49.4	260	4	US-10-437-963-138332	Sequence 138332,
33	41	49.4	412	5	US-10-739-930-10833	Sequence 10833, A
34	41	49.4	470	4	US-10-332-795-3	Sequence 3, Appli
35	41	49.4	470	5	US-10-626-832-8	Sequence 8, Appli
36	41	49.4	740	4	US-10-408-765A-2463	Sequence 2463, Ap
37	41	49.4	835	3	US-09-790-289-3	Sequence 3, Appli
38	41	49.4	835	4	US-10-468-582-3	Sequence 3, Appli
39	41	49.4	835	6	US-11-028-971-3	Sequence 3, Appli
40	41	49.4	1216	5	US-10-450-763-40113	Sequence 40113, A
41	41	49.4	1479	6	US-11-097-143-39600	Sequence 39600, A
42	40.5	48.8	52	4	US-10-029-386-29128	Sequence 29128, A
43	40	48.2	50	4	US-10-425-115-354560	Sequence 354560,
44	40	48.2	78	4	US-10-029-386-32475	Sequence 32475, A
45	40	48.2	80	4	US-10-424-599-204801	Sequence 204801,

ALIGNMENTS

RESULT 1
US-09-819-144A-8
; Sequence 8, Application US/09819144A
; Publication No. US20010016649A1
; GENERAL INFORMATION:
; APPLICANT: Terek, Richard M.
; TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES
; FILE REFERENCE: 21486-021DIV
; CURRENT APPLICATION NUMBER: US/09/819,144A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-819-144A-8

Query Match 100.0%; Score 83; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRQTLSHGSSSPARAC 16
||||||||||||||||
Db 1 RRQTLSHGSSSPARAC 16

RESULT 2
US-09-819-144A-2
; Sequence 2, Application US/09819144A
; Publication No. US20010016649A1
; GENERAL INFORMATION:
; APPLICANT: Terek, Richard M.
; TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES
; FILE REFERENCE: 21486-021DIV
; CURRENT APPLICATION NUMBER: US/09/819,144A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-819-144A-2

Query Match 89.2%; Score 74; DB 3; Length 52;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRQTLSHGSSSPARA 15

```
                ||||||||||||||||
Db          24 RRQTLSHGSSSPARA 38


RESULT 3
US-10-437-963-145735
; Sequence 145735, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 145735
;   LENGTH: 415
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(415)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_46427C.1.pep
US-10-437-963-145735

  Query Match             61.4%;  Score 51;  DB 4;  Length 415;
  Best Local Similarity   62.5%;  Pred. No. 6.6;
  Matches   10;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARAC 16
              || |   | ||||| |
Db        205 RRSTACRGGSSPARRC 220


RESULT 4
US-10-437-963-145733
; Sequence 145733, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 145733
;   LENGTH: 362
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(362)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_46425C.1.pep
US-10-437-963-145733

  Query Match             54.2%;  Score 45;  DB 4;  Length 362;
  Best Local Similarity   62.5%;  Pred. No. 57;
  Matches   10;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARAC 16
              || | |   ||||| |
Db        204 RRCTPSRAGSSPARRC 219


RESULT 5
US-10-085-108-26
; Sequence 26, Application US/10085108
; Publication No. US20020176865A1
;   GENERAL INFORMATION:
;        APPLICANT: LUCAS, Sophie; BOON-FALLEUR, Thierry
;        TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES CODING
;                            FOR
;                            TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C
;                            MAGE-B FAMILIES AND USES THEREOF
;        NUMBER OF SEQUENCES: 26
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fulbright & Jaworski L.L.P.
;             STREET: 666 Fifth Avenue
;             CITY: New York City
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10103
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
;             COMPUTER: IBM PS/2
;             OPERATING SYSTEM: PC-DOS
;             SOFTWARE: Wordperfect
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/085,108
;             FILING DATE: 01-Mar-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/501,104
;             FILING DATE: 09-Feb-2000
;             APPLICATION NUMBER: 09/468,433
;             FILING DATE: December 17, 1999
;             APPLICATION NUMBER: 09/066,281
;             FILING DATE: April 24, 1998
;             APPLICATION NUMBER: 08/845,528
;             FILING DATE: April 25, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mary Anne Schofield
;             REGISTRATION NUMBER: 36,669
;             REFERENCE/DOCKET NUMBER: LUD 5611.1 JEL/MAS
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 318-3100
;             TELEFAX: (212) 318-3400
;   INFORMATION FOR SEQ ID NO: 26:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 407
;             TYPE: amino acid
;             STRANDEDNESS: single-stranded
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-10-085-108-26

  Query Match             54.2%;  Score 45;  DB 4;  Length 407;
  Best Local Similarity   60.0%;  Pred. No. 64;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              :|  || ||| :|||
Db         34 KQEESHSSSSSSRAC 48


RESULT 6
```

```
US-10-108-260A-4444
; Sequence 4444, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20040005560A1el full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/10/108,260A
;  CURRENT FILING DATE:  2002-03-27
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4444
;   LENGTH: 407
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-108-260A-4444

  Query Match             54.2%;  Score 45;  DB 4;  Length 407;
  Best Local Similarity   60.0%;  Pred. No. 64;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              :|  || ||| :|||
Db         34 KQEESHSSSSSSRAC 48


RESULT 7
US-10-773-236-320
; Sequence 320, Application US/10773236
; Publication No. US20050208602A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et. al.
;  TITLE OF INVENTION: 89 Human Secreted Proteins
;  FILE REFERENCE: PS751P1
;  CURRENT APPLICATION NUMBER: US/10/773,236
;  CURRENT FILING DATE:  2004-02-09
;  PRIOR APPLICATION NUMBER: 60/311,085
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: 60/325,209
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: PCT/US02/25107
;  PRIOR FILING DATE: 2002-08-08
;  PRIOR APPLICATION NUMBER: 60/330,629
;  PRIOR FILING DATE: 2001-10-26
;  PRIOR APPLICATION NUMBER: PCT/US02/33985
;  PRIOR FILING DATE: 2002-10-24
;  PRIOR APPLICATION NUMBER: 60/331,046
;  PRIOR FILING DATE: 2001-11-07
;  PRIOR APPLICATION NUMBER: PCT/US02/35606
;  PRIOR FILING DATE: 2002-11-06
;  PRIOR APPLICATION NUMBER: 60/358,554
;  PRIOR FILING DATE: 2002-02-22
;  PRIOR APPLICATION NUMBER: PCT/US03/04819
;  PRIOR FILING DATE: 2003-02-20
;  PRIOR APPLICATION NUMBER: 60/358,714
;  PRIOR FILING DATE: 2002-02-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 396
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 320
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-773-236-320

  Query Match             53.0%;  Score 44;  DB 5;  Length 145;
  Best Local Similarity   56.2%;  Pred. No. 32;
  Matches    9;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARAC 16
              |   |||| | |  ||
Db         36 RFPVLSHGLSGPPHAC 51


RESULT 8
US-10-425-115-213570
; Sequence 213570, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 213570
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_126374C.1.pep
US-10-425-115-213570

  Query Match             53.0%;  Score 44;  DB 4;  Length 168;
  Best Local Similarity   61.5%;  Pred. No. 38;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPAR 14
              |  :||||: |||
Db        122 RMIVSHGSAGPAR 134


RESULT 9
US-10-437-963-140084
; Sequence 140084, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 140084
;   LENGTH: 1563
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_41315C.1.pep
US-10-437-963-140084

  Query Match             53.0%;  Score 44;  DB 4;  Length 1563;
  Best Local Similarity   50.0%;  Pred. No. 3.8e+02;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARAC 16
              ||: ::||| ||   |
Db        185 RRRLVNHGSRSPIVEC 200


RESULT 10
US-10-437-963-144158
; Sequence 144158, Application US/10437963
; Publication No. US20040123343A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 144158
;   LENGTH: 138
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_449C.1.pep
US-10-437-963-144158

  Query Match             51.8%;  Score 43;  DB 4;  Length 138;
  Best Local Similarity   63.6%;  Pred. No. 45;
  Matches    7;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSP 12
              | |::||:|||
Db        117 RDTIAHGNSSP 127


RESULT 11
US-10-425-115-199021
; Sequence 199021, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 199021
;   LENGTH: 198
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_113084C.1.pep
US-10-425-115-199021

  Query Match             51.8%;  Score 43;  DB 4;  Length 198;
  Best Local Similarity   53.8%;  Pred. No. 65;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 TLSHGSSSPARAC 16
              ::|||: ||  ||
Db        179 SISHGARSPLSAC 191


RESULT 12
US-10-437-963-143718
; Sequence 143718, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 143718
;   LENGTH: 378
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(378)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_445C.1.pep
US-10-437-963-143718

  Query Match             51.8%;  Score 43;  DB 4;  Length 378;
  Best Local Similarity   63.6%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSP 12
              | |::||:|||
Db        357 RDTIAHGNSSP 367


RESULT 13
US-10-437-963-134466
; Sequence 134466, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 134466
;   LENGTH: 711
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_36237C.1.pep
US-10-437-963-134466

  Query Match             51.8%;  Score 43;  DB 4;  Length 711;
  Best Local Similarity   64.3%;  Pred. No. 2.5e+02;
  Matches    9;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARA 15
              |   ||||| | ||
Db        422 RHLASHGSSMPRRA 435


RESULT 14
US-10-437-963-120904
; Sequence 120904, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
```

```
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 120904
;   LENGTH: 735
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_23981C.1.pep
US-10-437-963-120904

  Query Match             51.8%;  Score 43;  DB 4;  Length 735;
  Best Local Similarity   64.3%;  Pred. No. 2.5e+02;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARA 15
              || |:|   |||||
Db        614 RQLLAHTPDSPARA 627


RESULT 15
US-10-489-425-66
; Sequence 66, Application US/10489425
; Publication No. US20040248791A1
; GENERAL INFORMATION:
;  APPLICANT: Syngenta Participations AG
;  APPLICANT:  Spana, Eric
;  APPLICANT:  Kamdar, Kim
;  APPLICANT:  Stam, Lynn
;  APPLICANT:  Valentine, Scott
;  APPLICANT:  Griswald, Charles M
;  TITLE OF INVENTION: Insect G Protein-Coupled Receptor Genes and Uses thereof
;  FILE REFERENCE: 60134WOPCT
;  CURRENT APPLICATION NUMBER: US/10/489,425
;  CURRENT FILING DATE:  2004-03-08
;  PRIOR APPLICATION NUMBER: 60/341,512
;  PRIOR FILING DATE: 2001-12-18
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
;   LENGTH: 854
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-10-489-425-66

  Query Match             51.8%;  Score 43;  DB 5;  Length 854;
  Best Local Similarity   66.7%;  Pred. No. 3e+02;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 LSHGSSSPARAC 16
              | | ||||| :|
Db        603 LRHSSSSPASSC 614


Search completed: April 27, 2006, 15:40:42
Job time : 57 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 27, 2006, 15:36:55 ; Search time 8 Seconds
(without alignments)
90.956 Million cell updates/sec

Title: US-09-819-144A-8
Perfect score: 83
Sequence: 1 RRQTLSHGSSSPARAC 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 232119

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /SIDS5/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /SIDS5/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /SIDS5/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /SIDS5/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /SIDS5/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /SIDS5/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /SIDS5/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /SIDS5/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	53.0	553	7	US-11-188-298-12028	Sequence 12028, A
2	42	50.6	142	7	US-11-096-568A-21109	Sequence 21109, A
3	40	48.2	462	6	US-10-330-773-581	Sequence 581, App
4	40	48.2	537	7	US-11-037-243-104	Sequence 104, App
5	39	47.0	144	7	US-11-096-568A-9841	Sequence 9841, Ap
6	39	47.0	167	7	US-11-096-568A-9840	Sequence 9840, Ap
7	39	47.0	195	7	US-11-096-568A-9839	Sequence 9839, Ap
8	39	47.0	844	7	US-11-127-877-40	Sequence 40, Appl
9	39	47.0	1620	6	US-10-453-372-868	Sequence 868, App
10	38.5	46.4	985	6	US-10-216-161A-211	Sequence 211, App
11	38	45.8	118	7	US-11-096-568A-3538	Sequence 3538, Ap
12	38	45.8	149	7	US-11-072-512-3179	Sequence 3179, Ap
13	38	45.8	153	7	US-11-096-568A-21	Sequence 21, Appl
14	38	45.8	163	7	US-11-096-568A-20	Sequence 20, Appl
15	38	45.8	184	7	US-11-096-568A-19	Sequence 19, Appl
16	38	45.8	189	7	US-11-188-298-21058	Sequence 21058, A
17	38	45.8	199	6	US-10-467-657-4532	Sequence 4532, Ap
18	38	45.8	244	7	US-11-096-568A-4711	Sequence 4711, Ap
19	38	45.8	289	7	US-11-096-568A-4710	Sequence 4710, Ap
20	38	45.8	297	7	US-11-096-568A-24418	Sequence 24418, A
21	38	45.8	298	7	US-11-096-568A-4709	Sequence 4709, Ap
22	38	45.8	414	7	US-11-096-568A-19739	Sequence 19739, A
23	38	45.8	2197	7	US-11-075-185-8	Sequence 8, Appli
24	37	44.6	142	7	US-11-072-512-2706	Sequence 2706, Ap
25	37	44.6	168	6	US-10-467-657-5924	Sequence 5924, Ap
26	37	44.6	198	6	US-10-821-234-1448	Sequence 1448, Ap
27	37	44.6	332	7	US-11-045-004-304	Sequence 304, App
28	37	44.6	489	7	US-11-188-298-11028	Sequence 11028, A
29	37	44.6	574	7	US-11-188-298-2870	Sequence 2870, Ap
30	36	43.4	127	6	US-10-667-295-167	Sequence 167, App
31	36	43.4	196	6	US-10-467-657-8288	Sequence 8288, Ap
32	36	43.4	196	7	US-11-188-298-2795	Sequence 2795, Ap
33	36	43.4	196	7	US-11-188-298-6465	Sequence 6465, Ap
34	36	43.4	201	7	US-11-096-568A-21390	Sequence 21390, A
35	36	43.4	398	7	US-11-096-568A-2519	Sequence 2519, Ap
36	36	43.4	421	7	US-11-045-004-2837	Sequence 2837, Ap
37	36	43.4	430	7	US-11-072-512-2075	Sequence 2075, Ap
38	36	43.4	440	7	US-11-087-099-8921	Sequence 8921, Ap
39	36	43.4	453	7	US-11-096-568A-2518	Sequence 2518, Ap
40	36	43.4	474	7	US-11-188-298-4466	Sequence 4466, Ap
41	36	43.4	474	7	US-11-188-298-12062	Sequence 12062, A
42	36	43.4	487	7	US-11-096-568A-2517	Sequence 2517, Ap
43	36	43.4	552	7	US-11-294-997-38	Sequence 38, Appl
44	36	43.4	678	7	US-11-188-298-11543	Sequence 11543, A
45	36	43.4	809	7	US-11-091-018-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-11-188-298-12028
; Sequence 12028, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 12028
; LENGTH: 553
; TYPE: PRT
; ORGANISM: ASPERGILLUS NIDULANS FGSC A4
US-11-188-298-12028

Query Match 53.0%; Score 44; DB 7; Length 553;
Best Local Similarity 69.2%; Pred. No. 10;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy          3 QTLSHGSSSPARA 15
              ||  |||||: |
Db        491 QTTVHGSSSPSSA 503
```

RESULT 2
US-11-096-568A-21109
; Sequence 21109, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21109
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(142)
; OTHER INFORMATION: Ceres Seq. ID no. 12401370

US-11-096-568A-21109

Query Match 50.6%; Score 42; DB 7; Length 142;
Best Local Similarity 60.0%; Pred. No. 5.6;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 RRQTLSHGSSSPARA 15
              |  ||: | |:||||
Db        122 RSPTLASGRSAPARA 136
```

RESULT 3
US-10-330-773-581
; Sequence 581, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 581
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-330-773-581

Query Match 48.2%; Score 40; DB 6; Length 462;
Best Local Similarity 53.3%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```
Qy          2 RQTLSHGSSSPARAC 16
              |  |  || ||:| |
Db        388 RAPLEEGSLSPSREC 402
```

RESULT 4
US-11-037-243-104
; Sequence 104, Application US/11037243
; Publication No. US20050287546A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/11/037,243
; CURRENT FILING DATE: 2005-05-26
; PRIOR APPLICATION NUMBER: US/09/888,615
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-037-243-104

Query Match 48.2%; Score 40; DB 7; Length 537;
Best Local Similarity 77.8%; Pred. No. 48;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          6 SHGSSSPAR 14
              |||::||||
Db          5 SHGNASPAR 13
```

RESULT 5
US-11-096-568A-9841
; Sequence 9841, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 9841
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(144)
; OTHER INFORMATION: Ceres Seq. ID no. 12625408
US-11-096-568A-9841

Query Match 47.0%; Score 39; DB 7; Length 144;
Best Local Similarity 72.7%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          5 LSHGSSSPARA 15
              ||||:|||  |
Db        103 LSHGASSPPGA 113
```

RESULT 6
US-11-096-568A-9840
; Sequence 9840, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 9840
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(167)
; OTHER INFORMATION: Ceres Seq. ID no. 12625407
US-11-096-568A-9840

Query Match 47.0%; Score 39; DB 7; Length 167;
Best Local Similarity 72.7%; Pred. No. 21;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          5 LSHGSSSPARA 15
              ||||:|||  |
Db        126 LSHGASSPPGA 136
```

RESULT 7
US-11-096-568A-9839
; Sequence 9839, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2

; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 9839
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(195)
; OTHER INFORMATION: Ceres Seq. ID no. 12625406
US-11-096-568A-9839

Query Match 47.0%; Score 39; DB 7; Length 195;
Best Local Similarity 72.7%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 LSHGSSSPARA 15
||||:||| |
Db 154 LSHGASSPPGA 164

RESULT 8
US-11-127-877-40
; Sequence 40, Application US/11127877
; Publication No. US20050287565A1
; GENERAL INFORMATION:
; APPLICANT: Merchiers, Pascal G.
; APPLICANT: Hoffmann, Marcel
; APPLICANT: Spittaels, Koenraad F. F.
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Methods, Compositions and Compound Assays For Inhibiting
; TITLE OF INVENTION: Amyloid-Beta Protein Production
; FILE REFERENCE: P27,800-B USA
; CURRENT APPLICATION NUMBER: US/11/127,877
; CURRENT FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: 60/570,352
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/603,948
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 590
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 40
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-127-877-40

Query Match 47.0%; Score 39; DB 7; Length 844;
Best Local Similarity 88.9%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LSHGSSSPA 13
||:||||||
Db 148 LSYGSSSPA 156

RESULT 9
US-10-453-372-868
; Sequence 868, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 868
; LENGTH: 1620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-868

Query Match 47.0%; Score 39; DB 6; Length 1620;
Best Local Similarity 64.3%; Pred. No. 2.2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RRQTLSHGSSSPAR 14
|| || ||| | |
Db 36 RRTKLSRGSSVPPR 49

RESULT 10
US-10-216-161A-211
; Sequence 211, Application US/10216161A
; Publication No. US20060078964A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C91
; CURRENT APPLICATION NUMBER: US/10/216,161A
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: PCT/US00/04341
; PRIOR FILING DATE: 2000-02-18

```
;  PRIOR APPLICATION NUMBER: PCT/US99/05028
;  PRIOR FILING DATE: 1999-03-08
;  PRIOR APPLICATION NUMBER: US 09/380,138
;  PRIOR FILING DATE: 1999-08-25
;  PRIOR APPLICATION NUMBER: 60/126,773
;  PRIOR FILING DATE: 1999-03-29
;  PRIOR APPLICATION NUMBER: 60/081,955
;  PRIOR FILING DATE: 1998-04-15
;  NUMBER OF SEQ ID NOS: 624
; SEQ ID NO 211
;   LENGTH: 985
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-216-161A-211

  Query Match             46.4%;  Score 38.5;  DB 6;  Length 985;
  Best Local Similarity   36.0%;  Pred. No. 1.6e+02;
  Matches    9;  Conservative    3;  Mismatches    4;  Indels    9;  Gaps    1;

Qy          1 RRQTLSHGSSSP---------ARAC 16
              ::| | | :|||           |||
Db        625 KKQELQHANSSPLLRGSHSLELRAC 649


RESULT 11
US-11-096-568A-3538
; Sequence 3538, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3538
;   LENGTH: 118
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(118)
;   OTHER INFORMATION: Ceres Seq. ID no. 12611234
US-11-096-568A-3538

  Query Match             45.8%;  Score 38;  DB 7;  Length 118;
  Best Local Similarity   50.0%;  Pred. No. 22;
  Matches    7;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPAR 14
              ||:  :||   |||
Db        102 RRRAAAHGRHRPAR 115


RESULT 12
US-11-072-512-3179
; Sequence 3179, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3179
;   LENGTH: 149
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-3179

  Query Match             45.8%;  Score 38;  DB 7;  Length 149;
  Best Local Similarity   50.0%;  Pred. No. 28;
  Matches    8;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARAC 16
              |:|  |    || |||
Db          9 RKQPSSQKGPSPPRAC 24


RESULT 13
US-11-096-568A-21
; Sequence 21, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21
;   LENGTH: 153
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(153)
;   OTHER INFORMATION: Ceres Seq. ID no. 13586527
US-11-096-568A-21

  Query Match             45.8%;  Score 38;  DB 7;  Length 153;
  Best Local Similarity   42.9%;  Pred. No. 29;
  Matches    6;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPAR 14
              | :   |||::|:|
Db         49 RERAADHGSAAPSR 62


RESULT 14
US-11-096-568A-20
; Sequence 20, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
```

; SEQ ID NO 20
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(163)
; OTHER INFORMATION: Ceres Seq. ID no. 13586526
US-11-096-568A-20

Query Match 45.8%; Score 38; DB 7; Length 163;
Best Local Similarity 42.9%; Pred. No. 31;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RRQTLSHGSSSPAR 14
| : ||| ::|:|
Db 59 RERAADHGSAAPSR 72

RESULT 15
US-11-096-568A-19
; Sequence 19, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(184)
; OTHER INFORMATION: Ceres Seq. ID no. 13586525
US-11-096-568A-19

Query Match 45.8%; Score 38; DB 7; Length 184;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RRQTLSHGSSSPAR 14
| : ||| ::|:|
Db 80 RERAADHGSAAPSR 93

Search completed: April 27, 2006, 15:41:22
Job time : 9 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 27, 2006, 15:15:05 ; Search time 11.2941 Seconds
(without alignments)
136.307 Million cell updates/sec

Title: US-09-819-144A-8
Perfect score: 83
Sequence: 1 RRQTLSHGSSSPARAC 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	51.8	284	2	E84616	probable bHLH tran
2	43	51.8	624	2	T16452	hypothetical prote
3	42	50.6	437	2	S15144	hypothetical prote
4	41.5	50.0	667	2	T46094	hypothetical prote
5	41	49.4	285	2	T29490	hypothetical prote
6	41	49.4	437	2	C35147	integrase homolog
7	41	49.4	470	1	B42746	transcription acti
8	41	49.4	470	1	WZBEA4	transcription acti
9	40	48.2	315	2	C70810	hypothetical prote
10	40	48.2	330	2	G69169	hypothetical prote
11	40	48.2	527	2	S44081	11S globulin - oat
12	40	48.2	593	1	QQBEG6	HWLF1 protein - hu
13	40	48.2	639	2	C95342	nitrous-oxide redu
14	40	48.2	848	2	E71404	hypothetical prote
15	40	48.2	1266	2	I59314	isoleucine-tRNA li
16	40	48.2	1621	2	S62356	TRP-185 protein -
17	39	47.0	108	2	F72469	hypothetical prote
18	39	47.0	168	2	AF3294	peptidoglycan-asso
19	39	47.0	168	2	I40346	omp16 protein - Br
20	39	47.0	177	2	AE3013	omp16 protein [imp
21	39	47.0	177	2	B98271	omp16 protein [imp
22	39	47.0	187	2	S21466	nfxB protein - Pse
23	39	47.0	187	2	B83070	transcription regu
24	39	47.0	209	2	A71800	hypothetical prote
25	39	47.0	248	2	H96831	hypothetical prote
26	39	47.0	513	2	F85095	hypothetical prote
27	39	47.0	571	2	AE3281	hypothetical cytos
28	39	47.0	632	2	H84350	oligopeptidase [im
29	39	47.0	653	2	T01274	hypothetical prote
30	39	47.0	799	2	D85436	MAP3K-like protein
31	39	47.0	960	2	JE0356	gamma-aminobutyric
32	39	47.0	1891	2	T13594	hypothetical prote
33	39	47.0	1920	2	T13893	gene hindsight pro
34	38.5	46.4	158	2	D70755	hypothetical prote
35	38.5	46.4	479	2	T48025	hypothetical prote
36	38	45.8	137	2	T41575	ubiquinol-cytochro
37	38	45.8	203	2	B70784	cytochrome-c oxida
38	38	45.8	245	2	S43293	FLT3/FLK2 ligand (
39	38	45.8	271	2	A96011	probable cell divi
40	38	45.8	355	2	C96651	protein T3P18.9 [i
41	38	45.8	441	2	H70632	hypothetical prote
42	38	45.8	536	2	A45409	atrial natriuretic
43	38	45.8	549	1	F69361	arginyl-tRNA synth
44	38	45.8	735	2	T47594	hypothetical prote
45	38	45.8	960	2	S44812	F44B9.6 protein -

ALIGNMENTS

RESULT 1
E84616
probable bHLH transcription factor [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: E84616
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84616
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-284 <STO>
A;Cross-references: UNIPROT:O82397; UNIPARC:UPI000009F4A2; GB:AE002093; NID:g3738090; PI
C;Genetics:
A;Gene: At2g22760
A;Map position: 2

Query Match 51.8%; Score 43; DB 2; Length 284;
Best Local Similarity 53.3%; Pred. No. 11;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 RRQTLSHGSSSPARA 15
              :|:| |||: ||  |
Db        103 KRKTCSHGTRSPVLA 117
```

RESULT 2
T16452
hypothetical protein F53B3.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16452
R;Miller, N.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid F53B3.
A;Reference number: Z18515
A;Accession: T16452
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-624 <MIL>
A;Cross-references: UNIPROT:Q20712; UNIPARC:UPI000007F0CD; EMBL:U40414; NID:g1065474; PI
C;Genetics:
A;Gene: CESP:F53B3.3
A;Introns: 28/1; 53/2; 128/3; 167/3; 233/3; 288/1; 325/2; 473/3; 503/3; 549/1; 593/1
C;Superfamily: Caenorhabditis elegans hypothetical protein F53B3.3

Query Match 51.8%; Score 43; DB 2; Length 624;
Best Local Similarity 68.8%; Pred. No. 25;

```
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    4;  Gaps    1;

Qy          1 RRQTLSHGS----SSP 12
              | |||||||    |||
Db        277 RSQTLSHGSYAEPSSP 292


RESULT 3
S15144
hypothetical protein 437 precursor - phage Pf1
C;Species: phage Pf1
C;Date: 18-Feb-1994 #sequence_revision 19-Jan-1996 #text_change 09-Jul-2004
C;Accession: S15144; S20700
R;Hill, D.F.; Short, N.J.; Perham, R.N.; Petersen, G.B.
J. Mol. Biol. 218, 349-364, 1991
A;Title: DNA sequence of the filamentous bacteriophage Pf1.
A;Reference number: S15140; MUID:91186399; PMID:2010913
A;Accession: S15144
A;Molecule type: DNA
A;Residues: 1-437 <JMO>
A;Cross-references: UNIPROT:P25129; UNIPARC:UPI0000127C9A; EMBL:X52107; NID:g14829; PIDN
A;Experimental source: ATCC 25102-B1
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-437/Product: hypothetical protein 437 #status predicted <MAT>

  Query Match             50.6%;  Score 42;  DB 2;  Length 437;
  Best Local Similarity   72.7%;  Pred. No. 26;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 SHGSSSPARAC 16
              | |||:|| ||
Db         43 SSGSSTPAEAC 53


RESULT 4
T46094
hypothetical protein T25B15.10 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46094
R;Alcaraz, J.P.; Clabault, G.; Cottet, A.; Mache, R.; Mewes, H.W.; Lemcke, K.; Mayer, K.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23021
A;Accession: T46094
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-667 <ALC>
A;Cross-references: UNIPROT:Q9FT58; UNIPARC:UPI00000A2F6B; EMBL:AL132972
A;Experimental source: cultivar Columbia; BAC clone T25B15
C;Genetics:
A;Map position: 3
A;Introns: 11/2; 52/3; 108/1; 161/1; 255/3
A;Note: T25B15.10
C;Superfamily: Arabidopsis thaliana hypothetical protein T25B15.10

  Query Match             50.0%;  Score 41.5;  DB 2;  Length 667;
  Best Local Similarity   47.4%;  Pred. No. 47;
  Matches    9;  Conservative    3;  Mismatches    2;  Indels    5;  Gaps    1;

Qy          3 QTLSHGS-----SSPARAC 16
              :||||||     :|| : |
Db        305 ETLSHGSPDNSVTSPDKCC 323


RESULT 5
T29490
hypothetical protein M01H9.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Dec-2004
C;Accession: T29490
R;Pauley, A.; Gattung, S.
submitted to the EMBL Data Library, September 1996
A;Description: The sequence of C. elegans cosmid M01H9.
A;Reference number: Z20626
A;Accession: T29490
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-285 <PAU>
A;Cross-references: UNIPROT:Q94281; UNIPARC:UPI0000077876; EMBL:U70853; PIDN:AAB09145.1;
A;Experimental source: strain Bristol N2; clone M01H9
C;Genetics:
A;Gene: CESP:M01H9.4
A;Map position: 4
A;Introns: 16/3; 44/1; 90/1; 199/1

  Query Match             49.4%;  Score 41;  DB 2;  Length 285;
  Best Local Similarity   66.7%;  Pred. No. 25;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSP 12
              |||||| |:: |
Db        101 RRQTLSAGAAGP 112


RESULT 6
C35147
integrase homolog - Saccharopolyspora erythraea
C;Species: Saccharopolyspora erythraea
C;Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 09-Jul-2004
C;Accession: C35147
R;Brown, D.P.; Idler, K.B.; Katz, L.
J. Bacteriol. 172, 1877-1888, 1990
A;Title: Characterization of the genetic elements required for site-specific integration
A;Reference number: A35147; MUID:90202705; PMID:2180909
A;Accession: C35147
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-437 <BRO>
A;Cross-references: UNIPROT:P22877; UNIPARC:UPI000012D701; GB:M35138; NID:g152671; PIDN:

  Query Match             49.4%;  Score 41;  DB 2;  Length 437;
  Best Local Similarity   57.1%;  Pred. No. 38;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPAR 14
              :|||  || | | |
Db        240 QRQTWKHGCSDPHR 253


RESULT 7
B42746
transcription activator - equine herpesvirus 1 (strain Kentucky A)
C;Species: equine herpesvirus 1
A;Note: host Equus caballus (domestic horse)
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C;Accession: B42746
R;Zhao, Y.; Holden, V.R.; Harty, R.N.; O'Callaghan, D.J.
J. Virol. 66, 5363-5372, 1992
A;Title: Identification and transcriptional analyses of the UL3 and UL4 genes of equine
A;Reference number: A42746; MUID:92365125; PMID:1323700
A;Accession: B42746
A;Molecule type: DNA
A;Residues: 1-470 <ZHA>
A;Cross-references: UNIPROT:Q05906; UNIPARC:UPI000012D21C; GB:S43139
C;Superfamily: varicella-zoster virus gene 4 protein
C;Keywords: transcription regulation; zinc finger
F;441-466/Region: zinc finger

  Query Match             49.4%;  Score 41;  DB 1;  Length 470;
  Best Local Similarity   46.7%;  Pred. No. 41;
  Matches    7;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARA 15
              ||   :|||::| |:
Db        184 RRGNANHGSNTPGRS 198
```

RESULT 8
WZBEA4
transcription activator - equine herpesvirus 1 (strain Ab4p)
C;Species: equine herpesvirus 1
A;Note: host Equus caballus (domestic horse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C;Accession: F36795
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank, March 1992
A;Description: The DNA sequence of equine herpesvirus-1.
A;Reference number: A36805
A;Accession: F36795
A;Molecule type: DNA
A;Residues: 1-470 <TEL>
A;Cross-references: UNIPROT:P28939; UNIPARC:UPI0000047353; GB:M86664; NID:g330791; PIDN:
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A;Title: The DNA sequence of equine herpesvirus-1.
A;Reference number: A41831; MUID:92295566; PMID:1318606
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
C;Genetics:
A;Gene: 5
C;Superfamily: varicella-zoster virus gene 4 protein
C;Keywords: transcription regulation

  Query Match 49.4%; Score 41; DB 1; Length 470;
  Best Local Similarity 46.7%; Pred. No. 41;
  Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RRQTLSHGSSSPARA 15
Db 184 RRGNANHGSNTPGRS 198

RESULT 9
C70810
hypothetical protein Rv0819 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: C70810
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70810
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-315 <COL>
A;Cross-references: UNIPROT:O53831; UNIPARC:UPI00000D119B; GB:AL022004; GB:AL123456; NID
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv0819
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0819

  Query Match 48.2%; Score 40; DB 2; Length 315;
  Best Local Similarity 53.3%; Pred. No. 41;
  Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RRQTLSHGSSSPARA 15
Db 106 RNQFWAHGTLDPARA 120

RESULT 10
G69169
hypothetical protein MTH529 - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: G69169
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: G69169
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-330 <MTH>
A;Cross-references: UNIPROT:O26629; UNIPARC:UPI0000062B1E; GB:AE000836; GB:AE000666; NID
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH529
A;Start codon: GTG
C;Superfamily: Methanobacterium thermoautotrophicum hypothetical protein MTH529

  Query Match 48.2%; Score 40; DB 2; Length 330;
  Best Local Similarity 80.0%; Pred. No. 43;
  Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 SHGSSSPARA 15
Db 253 SHGSCPPARA 262

RESULT 11
S44081
11S globulin - oat
C;Species: Avena sativa (oat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Oct-2004
C;Accession: S44081
R;Tanchak, M.A.; Giband, M.; Potier, B.; Schernthaner, J.P.; Dukiandjiev, S.; Altosaar,
submitted to the EMBL Data Library, August 1993
A;Description: Genomic and cDNA clones encoding 11S globulins in oats (Avena sativa L.)
A;Reference number: S44081
A;Accession: S44081
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-527 <TAN>
A;Cross-references: UNIPROT:Q38780; UNIPARC:UPI00000A441A; EMBL:X74741; NID:g472866; PID
C;Genetics:
A;Introns: 110/1; 201/3; 381/3

  Query Match 48.2%; Score 40; DB 2; Length 527;
  Best Local Similarity 61.5%; Pred. No. 68;
  Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 QTLSHGSSSPARA 15
Db 513 QDIEEGSSSPVRA 525

RESULT 12
QQBEG6
HWLF1 protein - human cytomegalovirus (strain AD169)
N;Alternate names: hypothetical protein US22
C;Species: human cytomegalovirus, human herpesvirus 5
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004
C;Accession: F27231; S09936
R;Weston, K.; Barrell, B.G.
J. Mol. Biol. 192, 177-208, 1986
A;Title: Sequence of the short unique region, short repeats, and part of the long repeat
A;Reference number: A92935; MUID:87169717; PMID:3031311
A;Accession: F27231
A;Molecule type: DNA
A;Residues: 1-593 <WES>
A;Cross-references: UNIPROT:P09722; UNIPARC:UPI0000137E39; EMBL:X04650; NID:g59801; PIDN
A;Experimental source: strain AD169
R;Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990

```
A;Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A;Reference number: S09749; MUID:90269039; PMID:2161319
A;Accession: S09936
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-593 <CHE>
A;Cross-references: UNIPARC:UPI0000137E39; EMBL:X17403; NID:g59591; PIDN:CAA35289.1; PID
A;Experimental source: strain AD169
A;Note: this sequence was submitted to the EMBL Data Library, December 1989
A;Note: this reading frame extends between two stop codons and does not begin with a sta
C;Genetics:
A;Gene: HWLF1
C;Superfamily: cytomegalovirus HHLF5 protein

  Query Match             48.2%;  Score 40;  DB 1;  Length 593;
  Best Local Similarity   50.0%;  Pred. No. 76;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARAC 16
              ||::| | | :  |||
Db         25 RRRSLGHLSPTARRAC 40


RESULT 13
C95342
nitrous-oxide reductase (EC 1.7.99.6) NosZ [imported] - Sinorhizobium meliloti (strain 1
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95342
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: C95342
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-639 <KUR>
A;Cross-references: UNIPROT:Q59746; UNIPARC:UPI00001303B4; GB:AE006469; PIDN:AAK65301.1;
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: nosZ
A;Genome: plasmid
C;Superfamily: nitrous-oxide reductase
C;Keywords: oxidoreductase

  Query Match             48.2%;  Score 40;  DB 2;  Length 639;
  Best Local Similarity   66.7%;  Pred. No. 82;
  Matches    8;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 TLSHGSSSPARA 15
              ||| |:::||||
Db         35 TLSGGTATPARA 46


RESULT 14
E71404
hypothetical protein dl3180w - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
C;Accession: E71404
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: E71404
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-848 <BEV>
A;Cross-references: UNIPROT:O23286; UNIPARC:UPI00000A424D; GB:Z97335; NID:g2244747; PID:
C;Genetics:
A;Map position: 4COP9-4G3845
C;Superfamily: Arabidopsis thaliana hypothetical protein dl3180w

  Query Match             48.2%;  Score 40;  DB 2;  Length 848;
  Best Local Similarity   63.6%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSP 12
              |:|| || | |
Db        525 RETLHHGKSQP 535


RESULT 15
I59314
isoleucine-tRNA ligase (EC 6.1.1.5) - human
C;Species: Homo sapiens (man)
C;Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 09-Jul-2004
C;Accession: I59314
R;Shiba, K.; Suzuki, N.; Shigesada, K.; Namba, Y.; Schimmel, P.; Noda, T.
Proc. Natl. Acad. Sci. U.S.A. 91, 7435-7439, 1994
A;Title: Human cytoplasmic isoleucyl-tRNA synthetase: selective divergence of the antico
A;Reference number: I59314; MUID:94329529; PMID:8052601
A;Accession: I59314
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1266 <RES>
A;Cross-references: UNIPROT:P41252; UNIPARC:UPI00001364F3; GB:D28473; NID:g551621; PIDN:
C;Genetics:
A;Gene: GDB:IARS; ILRS
A;Cross-references: GDB:384085; OMIM:600709
A;Map position: 9q21-9q21
C;Superfamily: isoleucine-tRNA ligase
C;Keywords: ligase

  Query Match             48.2%;  Score 40;  DB 2;  Length 1266;
  Best Local Similarity   61.5%;  Pred. No. 1.6e+02;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 TLSHGSSSPARAC 16
              ||: ||| |  ||
Db       1072 TLTRGSSLPGPAC 1084


Search completed: April 27, 2006, 15:20:34
Job time : 13.2941 secs
```

OM protein - protein search, using sw model

Run on: April 27, 2006, 15:09:59 ; Search time 69.1765 Seconds
(without alignments)
163.183 Million cell updates/sec

Title: US-09-819-144A-8
Perfect score: 83
Sequence: 1 RRQTLSHGSSSPARAC 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	50	60.2	562	2	Q4Q9F1_LEIMA	Q4q9f1 leishmania
2	50	60.2	682	2	Q4S297_TETNG	Q4s297 tetraodon n
3	48	57.8	391	2	Q9RHC5_BRAJA	Q9rhc5 bradyrhizob
4	47	56.6	409	2	Q984F2_RHILO	Q984f2 rhizobium l
5	47	56.6	521	2	Q5B5R3_EMENI	Q5b5r3 aspergillus
6	47	56.6	840	1	MCPH1_HYLLA	P61592 hylobates l
7	46	55.4	820	2	Q6DFK4_XENLA	Q6dfk4 xenopus lae
8	46	55.4	1282	2	Q4QJ11_LEIMA	Q4qj11 leishmania
9	46	55.4	1588	2	Q6BUR9_DEBHA	Q6bur9 debaryomyce
10	45	54.2	388	2	Q6RAQ3_9PRIM	Q6raq3 bunopithecu
11	45	54.2	407	1	MAGB6_HUMAN	Q8n7x4 homo sapien
12	45	54.2	407	2	Q6GS19_HUMAN	Q6gs19 homo sapien
13	45	54.2	578	2	Q5V888_HALMA	Q5v888 haloarcula
14	44	53.0	203	2	Q4UBZ7_THEAN	Q4ubz7 theileria a
15	44	53.0	396	2	Q9AGT3_RHIME	Q9agt3 rhizobium m
16	44	53.0	431	2	Q6N2D2_RHOPA	Q6n2d2 rhodopseudo
17	44	53.0	553	2	Q5BAZ3_EMENI	Q5baz3 aspergillus
18	44	53.0	700	2	Q7QQ15_GIALA	Q7qq15 giardia lam
19	44	53.0	1077	2	O97217_LEIMA	O97217 leishmania
20	43	51.8	133	2	Q93JB4_STRCO	Q93jb4 streptomyce
21	43	51.8	147	2	Q5MRI2_DROME	Q5mri2 drosophila
22	43	51.8	232	2	Q9L1M7_STRCO	Q9l1m7 streptomyce
23	43	51.8	284	2	O82397_ARATH	O82397 arabidopsis
24	43	51.8	288	2	Q5MRH6_DROME	Q5mrh6 drosophila
25	43	51.8	288	2	Q5MRB0_DROME	Q5mrb0 drosophila
26	43	51.8	288	2	Q5MRB1_DROME	Q5mrb1 drosophila
27	43	51.8	288	2	Q5MRB3_DROME	Q5mrb3 drosophila
28	43	51.8	288	2	Q5MRB5_DROME	Q5mrb5 drosophila
29	43	51.8	288	2	Q5MRB6_DROME	Q5mrb6 drosophila
30	43	51.8	288	2	Q5MRC0_DROME	Q5mrc0 drosophila
31	43	51.8	288	2	Q5MRC1_DROME	Q5mrc1 drosophila
32	43	51.8	288	2	Q5MRC8_DROME	Q5mrc8 drosophila
33	43	51.8	288	2	Q5MRC9_DROME	Q5mrc9 drosophila
34	43	51.8	288	2	Q5MRD0_DROME	Q5mrd0 drosophila
35	43	51.8	288	2	Q5MRD1_DROME	Q5mrd1 drosophila
36	43	51.8	288	2	Q5MRD3_DROME	Q5mrd3 drosophila
37	43	51.8	288	2	Q5MRD6_DROME	Q5mrd6 drosophila
38	43	51.8	288	2	Q5MRD8_DROME	Q5mrd8 drosophila
39	43	51.8	288	2	Q5MRE4_DROME	Q5mre4 drosophila
40	43	51.8	288	2	Q5MRI1_DROME	Q5mri1 drosophila
41	43	51.8	288	2	Q5MRJ6_DROME	Q5mrj6 drosophila
42	43	51.8	288	2	Q5MRL7_DROME	Q5mrl7 drosophila
43	43	51.8	288	2	Q5MRL8_DROME	Q5mrl8 drosophila
44	43	51.8	288	2	Q5MRM0_DROME	Q5mrm0 drosophila
45	43	51.8	288	2	Q5MRQ4_DROME	Q5mrq4 drosophila

ALIGNMENTS

RESULT 1
Q4Q9F1_LEIMA
ID Q4Q9F1_LEIMA PRELIMINARY; PRT; 562 AA.
AC Q4Q9F1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=LmjF26.0400;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Fosker N., Harris D., Oliver K., O'Neil S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
DR EMBL; CT005265; CAJ04665.1; -; Genomic_DNA.
DR InterPro; IPR005123; 2OG-FeII_Oase.
DR Pfam; PF03171; 2OG-FeII_Oxy; 1.
KW Hypothetical protein; Iron; Oxidoreductase.
SQ SEQUENCE 562 AA; 60006 MW; DF991E7EC9686FF7 CRC64;

Query Match 60.2%; Score 50; DB 2; Length 562;
Best Local Similarity 64.3%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy          3 QTLSHGSSSPARAC 16
              |:||:|:| ||| |
Db        314 QSLSNGASPPARTC 327
```

RESULT 2
Q4S297_TETNG
ID Q4S297_TETNG PRELIMINARY; PRT; 682 AA.
AC Q4S297;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF14764, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00025195001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetradontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,

```
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014764; CAG05235.1; -; Genomic_DNA.
DR   InterPro; IPR001567; Pept_M3A_M3B.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01432; Peptidase_M3; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW   Hydrolase; Metalloprotease; Zinc.
FT   NON_TER     682    682
SQ   SEQUENCE   682 AA;  77979 MW;  F708954B6C07A113 CRC64;

  Query Match             60.2%;  Score 50;  DB 2;  Length 682;
  Best Local Similarity   61.5%;  Pred. No. 17;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 TLSHGSSSPARAC 16
              |: |||: |||:|
Db          2 TIQHGSAVPARSC 14


RESULT 3
Q9RHC5_BRAJA
ID   Q9RHC5_BRAJA PRELIMINARY;      PRT;   391 AA.
AC   Q9RHC5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative methyltransferase.
OS   Bradyrhizobium japonicum.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Bradyrhizobium.
OX   NCBI_TaxID=375;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=USDA 110spc4;
RX   MEDLINE=98418479; PubMed=9747707; DOI=10.1007/s004380050801;
RA   Becker B.U., Kosch K., Parniske M., Muller P.;
RT   "Exopolysaccharide (EPS) synthesis in Bradyrhizobium japonicum:
RT   sequence, operon structure and mutational analysis of an exo gene
RT   cluster.";
RL   Mol. Gen. Genet. 259:161-171(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=USDA 110spc4;
RA   Becker B.U., Kosch K., Parniske M., Muller P.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF039306; AAF19662.1; -; Genomic_DNA.
DR   HSSP; Q9WZK7; 1O54.
DR   GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR000682; PCMT.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF01135; PCMT; 1.
KW   Methyltransferase; Transferase.
SQ   SEQUENCE   391 AA;  42535 MW;  9DA2D9EBD37FA7CC CRC64;

  Query Match             57.8%;  Score 48;  DB 2;  Length 391;
  Best Local Similarity   53.3%;  Pred. No. 19;
  Matches    8;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              |: |:||:: | |||
Db        123 RRYLTHGATMPRRAC 137


RESULT 4
Q984F2_RHILO
ID   Q984F2_RHILO PRELIMINARY;      PRT;   409 AA.
AC   Q984F2;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Mlr8031 protein.
GN   OrderedLocusNames=mlr8031;
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; BA000012; BAB53678.1; -; Genomic_DNA.
DR   GO; GO:0016747; F:transferase activity, transferring groups o. . .; IEA.
DR   InterPro; IPR002656; Acyl_transf_3.
DR   Pfam; PF01757; Acyl_transf_3; 1.
KW   Complete proteome.
SQ   SEQUENCE   409 AA;  45649 MW;  F0BB59561BFDB410 CRC64;

  Query Match             56.6%;  Score 47;  DB 2;  Length 409;
  Best Local Similarity   57.1%;  Pred. No. 29;
  Matches    8;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 QTLSHGSSSPARAC 16
              :|  ||:|:| |||
Db         11 RTQLHGNSNPCRAC 24


RESULT 5
Q5B5R3_EMENI
ID   Q5B5R3_EMENI PRELIMINARY;      PRT;   521 AA.
AC   Q5B5R3;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=AN4117.2;
OS   Aspergillus nidulans FGSC A4.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=227321;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=FGSC A4;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
```

```
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Genome Sequence of Aspergillus nidulans.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Obtusifoliol + 3 O(2) + 3 NADPH = 4-alpha-
CC       methyl-5-alpha-ergosta-8,14,24(28)-trien-3-beta-ol + formate + 3
CC       NADP(+) + 3 H(2)O.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AACD01000067; EAA59378.1; -; Genomic_DNA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   GO; GO:0016126; P:sterol biosynthesis; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002403; EP450IV.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00465; EP450IV.
KW   Heme; Hypothetical protein; Lipid synthesis; Membrane; Monooxygenase;
KW   Oxidoreductase; Steroid biosynthesis; Sterol biosynthesis.
SQ   SEQUENCE   521 AA;  57178 MW;  E19D8BAFE91CC0B6 CRC64;

  Query Match             56.6%;  Score 47;  DB 2;  Length 521;
  Best Local Similarity   69.2%;  Pred. No. 39;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPAR 14
              |||  |||:|| |
Db        411 RQTTPHGSTSPRR 423


RESULT 6
MCPH1_HYLLA
ID   MCPH1_HYLLA    STANDARD;      PRT;   840 AA.
AC   P61592;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Microcephalin.
GN   Name=MCPH1;
OS   Hylobates lar (Common gibbon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Hylobatidae; Hylobates.
OX   NCBI_TaxID=9580;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX   PubMed=15056607; DOI=10.1093/hmg/ddh126;
RA   Evans P.D., Anderson J.R., Vallender E.J., Choi S.S., Lahn B.T.;
RT   "Reconstructing the evolutionary history of microcephalin, a gene
RT   controlling human brain size.";
RL   Hum. Mol. Genet. 13:1139-1145(2004).
CC   -!- FUNCTION: Implicated in chromosome condensation and DNA damage
CC       induced cellular responses. May play a role in neurogenesis and
CC       regulation of the size of the cerebral cortex (By similarity).
CC   -!- SIMILARITY: Contains 3 BRCT domains.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AY553037; AAS91380.1; -; Genomic_DNA.
DR   EMBL; AY553024; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553025; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553026; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553027; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553028; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553029; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553030; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553031; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553032; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553033; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553034; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553035; AAS91380.1; JOINED; Genomic_DNA.
DR   EMBL; AY553036; AAS91380.1; JOINED; Genomic_DNA.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR002378; Brst_cancerI.
DR   Pfam; PF00533; BRCT; 3.
DR   PRINTS; PR00493; BRSTCANCERI.
DR   SMART; SM00292; BRCT; 3.
DR   PROSITE; PS50172; BRCT; 3.
KW   Repeat.
FT   DOMAIN        1     93       BRCT 1.
FT   DOMAIN      644    734       BRCT 2.
FT   DOMAIN      755    837       BRCT 3.
SQ   SEQUENCE   840 AA;  92887 MW;  51265B4274C8CB87 CRC64;

  Query Match             56.6%;  Score 47;  DB 1;  Length 840;
  Best Local Similarity   57.1%;  Pred. No. 67;
  Matches    8;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPAR 14
              :|| :|||| ||::
Db        355 KRQRVSHGSHSPSK 368


RESULT 7
Q6DFK4_XENLA
ID   Q6DFK4_XENLA PRELIMINARY;      PRT;   820 AA.
AC   Q6DFK4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC81344 protein.
GN   Name=MGC81344;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
```

```
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC076732; AAH76732.1; -; mRNA.
SQ   SEQUENCE   820 AA;  89973 MW;  9ACB8A0CB3A2726C CRC64;

  Query Match             55.4%;  Score 46;  DB 2;  Length 820;
  Best Local Similarity   71.4%;  Pred. No. 95;
  Matches   10;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARA 15
              ||| ||| | || |
Db        316 RQTESHGVSDPAAA 329


RESULT 8
Q4QJ11_LEIMA
ID   Q4QJ11_LEIMA PRELIMINARY;      PRT;  1282 AA.
AC   Q4QJ11;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=LmjF06.0670;
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Friedlin;
RA   Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA   Smith D., Collins M., Fosker N., Harris D., Oliver K., O'Neil S.,
RA   Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL   Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CT005245; CAJ02112.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   1282 AA;  130054 MW;  D8C5587E782D6BAB CRC64;

  Query Match             55.4%;  Score 46;  DB 2;  Length 1282;
  Best Local Similarity   53.3%;  Pred. No. 1.6e+02;
  Matches    8;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARA 15
              || :|:|||| | ::
Db        906 RRSSLTHGSSPPEKS 920


RESULT 9
Q6BUR9_DEBHA
ID   Q6BUR9_DEBHA PRELIMINARY;      PRT;  1588 AA.
AC   Q6BUR9;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Similarities with sp|P14284 Saccharomyces cerevisiae YPL167c
DE   REV3.
GN   OrderedLocusNames=DEHA0C09438g;
OS   Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX   NCBI_TaxID=4959;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=ATCC 36239 / CBS 767;
RX   PubMed=15229592; DOI=10.1038/nature02579;
RA   Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA   Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA   Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA   Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA   Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA   Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA   Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA   Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA   Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA   Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA   Swennen D., Tekaia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA   Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA   Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA   Wincker P., Souciet J.-L.;
RT   "Genome evolution in yeasts.";
RL   Nature 430:35-44(2004).
DR   EMBL; CR382135; CAG86117.1; -; Genomic_DNA.
DR   GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR   GO; GO:0000166; F:nucleotide binding; IEA.
DR   GO; GO:0006260; P:DNA replication; IEA.
DR   InterPro; IPR006172; DNA_pol_B.
DR   InterPro; IPR006133; DNA_pol_B_exo.
DR   InterPro; IPR006134; DNA_pol_B_region.
DR   InterPro; IPR012337; RNaseH_fold.
DR   Pfam; PF00136; DNA_pol_B; 1.
DR   Pfam; PF03104; DNA_pol_B_exo; 1.
DR   PRINTS; PR00106; DNAPOLB.
DR   SMART; SM00486; POLBc; 1.
DR   PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW   Complete proteome.
SQ   SEQUENCE   1588 AA;  184486 MW;  C8667DD8D9E60772 CRC64;

  Query Match             55.4%;  Score 46;  DB 2;  Length 1588;
  Best Local Similarity   66.7%;  Pred. No. 2e+02;
  Matches   10;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              :|||| ||| |||
Db        657 QQTLSDSSSSKIRAC 671


RESULT 10
Q6RAQ3_9PRIM
ID   Q6RAQ3_9PRIM PRELIMINARY;      PRT;   388 AA.
AC   Q6RAQ3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Microcephalin (Fragment).
GN   Name=Mcph1;
OS   Bunopithecus hoolock (Hoolock gibbon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Hylobatidae; Bunopithecus.
OX   NCBI_TaxID=61851;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15056608; DOI=10.1093/hmg/ddh127;
RA   Wang Y.Q., Su B.;
RT   "Molecular evolution of microcephalin, a gene determining human brain
RT   size.";
RL   Hum. Mol. Genet. 13:1131-1137(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
```

```
RA   Wang Y., Su B.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY505993; AAS87830.1; -; Genomic_DNA.
FT   NON_TER       1      1
FT   NON_TER     388    388
SQ   SEQUENCE   388 AA;  42086 MW;  9AFB20E9B10F8BAB CRC64;

  Query Match             54.2%;  Score 45;  DB 2;  Length 388;
  Best Local Similarity   57.1%;  Pred. No. 60;
  Matches    8;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPAR 14
              :|| :|||| || :
Db        131 KRQRVSHGSHSPPK 144


RESULT 11
MAGB6_HUMAN
ID   MAGB6_HUMAN    STANDARD;      PRT;   407 AA.
AC   Q8N7X4; Q9H219;
DT   29-MAR-2004 (Rel. 43, Created)
DT   29-MAR-2004 (Rel. 43, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Melanoma-associated antigen B6 (MAGE-B6 antigen).
GN   Name=MAGEB6;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Testis;
RX   PubMed=14702039; DOI=10.1038/ng1285;
RA   Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA   Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA   Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA   Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA   Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA   Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA   Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA   Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA   Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA   Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA   Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA   Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA   Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA   Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA   Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA   Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA   Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA   Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA   Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA   Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA   Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA   Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA   Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA   Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA   Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA   Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA   Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA   Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT   "Complete sequencing and characterization of 21,243 full-length human
RT   cDNAs.";
RL   Nat. Genet. 36:40-45(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE OF 49-177, AND TISSUE SPECIFICITY.
RC   TISSUE=Blood;
RX   MEDLINE=20321428; PubMed=10861452;
RA   Lucas S., De Plaen E., Boon T.;
RT   "MAGE-B5, MAGE-B6, MAGE-C2, and MAGE-C3: four new members of the MAGE
RT   family with tumor-specific expression.";
RL   Int. J. Cancer 87:55-60(2000).
CC   -!- TISSUE SPECIFICITY: Expressed in testis. Not expressed in other
CC       normal tissues, but is expressed in tumors of different
CC       histological origins.
CC   -!- SIMILARITY: Contains 1 MAGE domain.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AK097561; BAC05102.1; -; mRNA.
DR   EMBL; AF320514; AAG48624.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000176746; Homo sapiens.
DR   HGNC; HGNC:23796; MAGEB6.
DR   MIM; 300467; -.
DR   InterPro; IPR002190; MAGE.
DR   PANTHER; PTHR11736; MAGE; 1.
DR   Pfam; PF01454; MAGE; 1.
DR   PROSITE; PS50838; MAGE; 1.
KW   Antigen; Multigene family.
FT   DOMAIN      195    394       MAGE.
FT   COMPBIAS     38    164       Ser-rich.
FT   CONFLICT    172    172       R -> G (in Ref. 2).
SQ   SEQUENCE   407 AA;  44091 MW;  D674F426131244C0 CRC64;

  Query Match             54.2%;  Score 45;  DB 1;  Length 407;
  Best Local Similarity   60.0%;  Pred. No. 63;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              :|  || ||| :|||
Db         34 KQEESHSSSSSSRAC 48


RESULT 12
Q6GS19_HUMAN
ID   Q6GS19_HUMAN PRELIMINARY;      PRT;   407 AA.
AC   Q6GS19;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Melanoma antigen, family B, 6 protein.
GN   Name=MAGEB6;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Pooled tissue, and Testis;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
```

```
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Testis;
RA   Director MGC Project;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Pooled tissue;
RG   NIH MGC Project;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC067286; AAH67286.1; -; mRNA.
DR   EMBL; BC074920; AAH74920.1; -; mRNA.
DR   Ensembl; ENSG00000176746; Homo sapiens.
DR   InterPro; IPR002190; MAGE.
DR   Pfam; PF01454; MAGE; 1.
DR   PROSITE; PS50838; MAGE; 1.
SQ   SEQUENCE   407 AA;  43992 MW;  3674F422020355DA CRC64;

  Query Match             54.2%;  Score 45;  DB 2;  Length 407;
  Best Local Similarity   60.0%;  Pred. No. 63;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RQTLSHGSSSPARAC 16
              :|  || ||| :|||
Db         34 KQEESHSSSSSSRAC 48


RESULT 13
Q5V888_HALMA
ID   Q5V888_HALMA PRELIMINARY;      PRT;   578 AA.
AC   Q5V888;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=pNG1039;
OS   Haloarcula marismortui (Halobacterium marismortui).
OG   Plasmid pNG100.
OC   Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC   Halobacteriaceae; Haloarcula.
OX   NCBI_TaxID=2238;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 43049;
RX   PubMed=15520287; DOI=10.1101/gr.2700304;
RA   Baliga N.S., Bonneau R., Facciotti M.T., Pan M., Glusman G.,
RA   Deutsch E.W., Shannon P., Chiu Y., Weng R.S., Gan R.R., Hung P.,
RA   Date S.V., Marcotte E., Hood L., Ng W.V.;
RT   "Genome sequence of Haloarcula marismortui: a halophilic archaeon from
RT   the Dead Sea.";
RL   Genome Res. 14:2221-2234(2004).
DR   EMBL; AY596290; AAV44292.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein; Plasmid.
SQ   SEQUENCE   578 AA;  63947 MW;  8D6BC42A493613E7 CRC64;

  Query Match             54.2%;  Score 45;  DB 2;  Length 578;
  Best Local Similarity   56.2%;  Pred. No. 94;
  Matches    9;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARAC 16
              ||   |:|||    |||
Db        208 RRSRLTHGSRRNRRAC 223


RESULT 14
Q4UBZ7_THEAN
ID   Q4UBZ7_THEAN PRELIMINARY;      PRT;   203 AA.
AC   Q4UBZ7;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=TA04580;
OS   Theileria annulata.
OC   Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileriidae;
OC   Theileria.
OX   NCBI_TaxID=5874;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Ankara isolate clone C9;
RA   Pain A., Renauld H., Murphy L., Harris D.A., Quail M.A., Berriman M.,
RA   Hall N., Barrell B.G.;
RT   "The chromosome 3 genome sequence of Theileria annulata.";
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR940352; CAI75654.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   203 AA;  22921 MW;  BB2593EA74D7A982 CRC64;

  Query Match             53.0%;  Score 44;  DB 2;  Length 203;
  Best Local Similarity   50.0%;  Pred. No. 42;
  Matches    7;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPAR 14
              | :|:|||  ||::
Db         69 RSKTMSHGGESPSK 82


RESULT 15
Q9AGT3_RHIME
ID   Q9AGT3_RHIME PRELIMINARY;      PRT;   396 AA.
AC   Q9AGT3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Rep protein.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OG   Plasmid pRm1132f.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21172893; PubMed=11274136;
RX   DOI=10.1128/JB.183.8.2704-2708.2001;
RA   Barran L.R., Ritchot N., Bromfield E.S.;
RT   "Sinorhizobium meliloti plasmid pRm1132f replicates by a rolling-
RT   circle mechanism.";
RL   J. Bacteriol. 183:2704-2708(2001).
DR   EMBL; AF327371; AAG59887.1; -; Genomic_DNA.
KW   Plasmid.
SQ   SEQUENCE   396 AA;  42948 MW;  89EA71F998579FE7 CRC64;

  Query Match             53.0%;  Score 44;  DB 2;  Length 396;
  Best Local Similarity   56.2%;  Pred. No. 90;
  Matches    9;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 RRQTLSHGSSSPARAC 16
              ||:   ||  | ||||
Db        346 RRRGRPHGDGSGARAC 361


Search completed: April 27, 2006, 15:19:42
Job time : 71.1765 secs
```